







```
FT CARBOHYD 803 803 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 843 AA; 95648 MW; C69DFD971C918B71 CRC64;

Query Match 82.6%; Score 2869.5; DB 1; Length 843;
Best Local Similarity 83.5%; Pred. No. 1.4e-219;
Matches 541; Conservative 35; Mismatches 47; Indels 25; Gaps 6;

Qy 3 LWVTYVYGVVPWKAEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 62
Dy 33 LWVTYVYGVVPWKAEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 92
Qy 63 NMWKNVVEQVHEDIISLWQSLKPCVKLTPLCVLTNCTDINTNTTNTTSLSIIVVWEQ 122
Dy 93 NMWKNVVEQVHEDIISLWQSLKPCVKLTPLCVLTNCTDINTNTTSSS---WET 146

Qy 123 RGKEMRNCSENIITSIRDKVQREYALFYKLDVDFIDNNKNTNTKYRLINCNTSVITQ 182
Dy 147 MEKGEIKNCSENIITSIRDKVQREYALFYKLDVDFIDNNKNTNTKYRLINCNTSVITQ 199
Qy 183 ACQVSVPEPIPIHYCTPTGFRALLKCNCKKFGTGTCTNVSTVQCTHGIRPVVSTOLLNG 242
Dy 200 ACQVSVPEPIPIHYCAPAGFAILKCNCKKFGTGTCTNVSTVQCTHGIRPVVSTOLLNG 259
Qy 243 SLAEEVIVRSENFNTNAKTIIVQLNVSVENCTRPNNHTRKRVLTGPRVWYTTGETLG 302
Dy 260 SLAEEVIVRSENFNTNAKTIIVQLNVSVENCTRPNNHTRKRVLTGPRVWYTTGETLG 319
Qy 303 NROAHNCNISRAQNNLTQIATTLRQFG-NKTIAPNQSSGGDPEIVMHSFNCGGSEFFY 361
Dy 320 DIRQAHNCNLSTQWENTLEQIAIKLQEGFGNKKTIIFNPSSGGDPEIVTHSFNCGGSEFFY 379
Qy 362 CNSQLNSANVTSNGTWSVTRKQKTGDIITPCRIKQIINWQVGVKAWYALPKGL 421
Dy 380 CNSQLP-----TWNDTRKLNNTGRNITPCRIKQIINWQVGVKAWYALPKGL 429
Qy 422 IRCSNITGLLITRDGGGENTTIFRPGGDMEDNWRSELYKVKVIEPLGVAPTAK 481
Dy 430 IRCSNITGLLITRDGGGKOTNGTIFRPGGDMEDNWRSELYKVKVIEPLGVAPTAK 489
Qy 482 RRVQREKRAVGMGAMFLGFLGAAGSTMGATSMALTVQARQLLSGIVQOQNLLRAIKA 541
Dy 490 RRVQREKRAVG-LGALFLGFLGAAGSTMGASITLTVQARQLLSGIVQOQNLLRAIEA 548
Qy 542 QOHLQLTVWGIKQOARILAVERYLKDQQLLGGWCGSKLICITVAPWNASWKNLTQ 601
Dy 549 QOHLQLTVWGIKQOARILAVERYLKDQQLLGGWCGSKLICITVAPWNASWKNLSNE 608
Qy 602 IWNNTWMDREIDNTYTHLYTTLIESQOQEKNOQELLQLDKWSL 649
Dy 609 IWDNMTWMDREIDNTYTHLYTTLIESQOQEKNOQELLQLDKWSL 656

RESULT 4
ENV_HV1W1 STANDARD; PRT; 856 AA.
ID ENV_HV1W1
AC P31872;
DT 01-JUL-1993 (Rel. 26, Created)
DI 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ1 isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=31678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86218077; PubMed=24233250;
RA Starch B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H.,
RA Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F.;
RT "Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS.";
```

```
Cell 45:637-648(1986).
-1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
WAS PERINATALLY INFECTED BY HER MOTHER.
PIR; A24774; VCLJ3W.
PDB; 1LB0; 04-DEC-02.
PDB; 1LCX; 04-DEC-02.
InterPro; IPR000328; Env GP41.
InterPro; IPR000777; GP120.
Pfam; PF00516; GP120; 1.
Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
3D-structure.
SIGNAL 1 29
CHAIN 30 510
CHAIN 511 856
DIULFID 53 73
DIULFID 118 205
DIULFID 125 196
DIULFID 130 152
DIULFID 218 247
DIULFID 228 239
DIULFID 296 330
DIULFID 376 444
DIULFID 383 417
CARBOHYD 87 87
CARBOHYD 134 134
CARBOHYD 140 140
CARBOHYD 151 151
CARBOHYD 155 155
CARBOHYD 183 183
CARBOHYD 197 197
CARBOHYD 234 234
CARBOHYD 241 241
CARBOHYD 262 262
CARBOHYD 276 276
CARBOHYD 289 289
CARBOHYD 295 295
CARBOHYD 331 331
CARBOHYD 338 338
CARBOHYD 354 354
CARBOHYD 360 360
CARBOHYD 390 390
CARBOHYD 394 394
CARBOHYD 404 404
CARBOHYD 447 447
CARBOHYD 459 459
CARBOHYD 611 611
CARBOHYD 616 616
CARBOHYD 625 625
CARBOHYD 637 637
SQ SEQUENCE 856 AA; 97526 MW; DB68D1E49C404DE9 CRC64;

Query Match 82.3%; Score 2858; DB 1; Length 856;
Best Local Similarity 82.8%; Pred. No. 1.2e-218;
Matches 539; Conservative 34; Mismatches 60; Indels 18; Gaps 5;

Qy 3 LWVTYVYGVVPWKAEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 62
Dy 33 LWVTYVYGVVPWKAEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 92
Qy 63 NMWKNVVEQVHEDIISLWQSLKPCVKLTPLCVLTNCTDINTNTTNTTSLSIIVVWEQ 122
Dy 93 NMWKNVVEQVHEDIISLWQSLKPCVKLTPLCVLTNCTDINTNTTNTTSLSIIVVWEQ 140
Qy 123 R---GKEMRNCSENIITSIRDKVQREYALFYKLDVDFIDNNKNTNTKYRLINCNTSV 179
Dy 141 KTIIGGEVNCSENIITSIRDKVQREYALFYKLDVDFIDNNKNTNTKYRLINCNTSV 200
Qy 180 ITQACPKVSEPIPIHYCTPTGFRALLKCNCKKFGTGTCTNVSTVQCTHGIRPVVSTOLL 239
Dy 201 ITQACPKVSEPIPIHYCAPAGFAILKCNCKKFGTGTCTNVSTVQCTHGIRPVVSTOLL 260
```

QY 240 LNSGLAEVEVIRSENFNTNAKIIIVQLAVSVINCTRPNNHTRKR-VTLGPGRVVYTTG 298  
 DB 261 LNSGLAEVEVIRSENFNTNAKIIIVQLAVSVINCTRPNNHTRKR-VTLGPGRVVYTTG 319  
 QY 299 EILGNTROAHCNISRAQWNTLOQIATTLREQFGKNTIAFNQSSGGDPPEIVMHSFNCGGE 358  
 DB 320 EIRGNTROAHCNISRAQWNTLOQIATTLREQFGKNTIAFNQSSGGDPPEIVMHSFNCGGE 379  
 QY 359 FFYCSNQLFNSAMVNTSGTWSVTRKQDGTDIITLPCRKIQIINRWQVVGKAMVALPI 418  
 DB 380 FFYCSNQLFNSAMVNTSGTWSVTRKQDGTDIITLPCRKIQIINRWQVVGKAMVALPI 438  
 QY 419 KGLIRCSSNITGLLLTRDGGENQTTIEFRPGGDMRDNRWSLYKYKVKIBPLGVAPT 478  
 DB 439 GGOIRCSSNITGLLLTRDGGENQTTIEFRPGGDMRDNRWSLYKYKVKIBPLGVAPT 498  
 QY 479 KAKRRVVQREKRAVGMGLFGLAAGTMTGATSMALTVQARQLLSGIVQOQNNLLRA 538  
 DB 499 KAKRRVVQREKRAVGMGLFGLAAGTMTGATSMALTVQARQLLSGIVQOQNNLLRA 558  
 QY 539 IKAQOHLQLTVWGIKQLQARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASNKT 598  
 DB 559 IEAQOHLQLTVWGIKQLQARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASNKT 618  
 QY 599 LDOIWNMTWMEWDREIDNTHLYIILIESQOQEKNOQELLODKWASL 649  
 DB 619 MDQIWNMTWMEWDREIDNTHLYIILIESQOQEKNOQELLODKWASL 669

RESULT 5  
 ENV\_HV1S3 STANDARD; PRT; 852 AA.

AC P19549;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Envelope polyprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90317906; PubMed=2370688;  
 RA York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;  
 RT "Human immunodeficiency virus type 1 cellular host range, replication, and cytopathicity are linked to the envelope region of the viral genome.";  
 RL J. Virol. 64:4016-4020(1990).

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M38427; AAA45067.1; --  
 DR PDB; 1MRQ; 11-DEC-02.  
 DR HIV; M38427; ENVSSF33.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 31 BY SIMILARITY.  
 FT CHAIN 32 506 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 507 852 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 73 BY SIMILARITY.  
 FT DISULFID 118 206 BY SIMILARITY.

FT DISULFID 125 197 BY SIMILARITY.  
 FT DISULFID 130 156 BY SIMILARITY.  
 FT DISULFID 219 248 BY SIMILARITY.  
 FT DISULFID 229 240 BY SIMILARITY.  
 FT DISULFID 297 331 BY SIMILARITY.  
 FT DISULFID 377 439 BY SIMILARITY.  
 FT DISULFID 384 412 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 242 242 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 391 391 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 621 621 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 633 633 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 812 812 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 852 AA; 56663 MW; 5E7BF8D23C9910D CRC64;

Query Match 82.0%; Score 2850; DB 1; Length 852;  
 Best Local Similarity 81.9%; Pred. No. 5,1e-218;  
 Matches 533; Conservative 51; Mismatches 47; Indels 20; Gaps 7;

QY 2 NLWTVYGVVWKEATTLTFCASDAKAYDTEAHNVWATHACVPTNPQEVVLNVNTE 61  
 DB 32 NLWTVYGVVWKEATTLTFCASDAKAYDTEAHNVWATHACVPTNPQEVVLNVNTE 91  
 QY 62 FNNWKNMVQEHEDIISLDQSLKPCVKLTPLCVTLNCTDL--NTNNTNTTSLIUV 119  
 DB 92 FNNWKNMVQEHEDIISLDQSLKPCVKLTPLCVTLNCTDLNATNNSGGTV--- 148  
 QY 120 WEQKGEMRNCNFNTTSIRDKVQREYALFYKLDVEPIDDNKNTTNTKRLINCTSV 179  
 DB 149 ---EKEIKNCNFNTTIRDKVQREYALFYKLDVVPIDDD-NT--NTSYRLHCNSSV 201  
 QY 180 ITQACPKVSEPIPHYCTPTGTFALLKNDKXGPGCTNVSTVQCTGIRPVSTOLL 239  
 DB 202 ITQTCPKVSEPIPHYCTPTGTFALLKNDKXGPGCTNVSTVQCTGIRPVSTOLL 261  
 QY 240 LNSGLAEVEVIRSENFNTNAKIIIVQLAVSVINCTRPNNHTRKRVTLPGRVWYTTGE 299  
 DB 262 LNSGLAEVEVIRSENFNTNAKIIIVQLAVSVINCTRPNNHTRKRVTLPGRVWYTTGE 321  
 QY 300 ILGNIRQAHCNISRAQWNTLOQIATTLREQFGKNTIAFNQSSGGDPPEIVMHSFNCGGE 359  
 DB 322 IIGDIRKAYCNISRAQWNTLOQIATTLREQFGKNTIAFNQSSGGDPPEIVMHSFNCGGE 391  
 QY 360 FYCNSQLFNSAMVNTSGTWSVTRKQDGTDIITLPCRKIQIINRWQVVGKAMVALPIK 419  
 DB 382 FYCNSQLFNSAMVNTSGTWSVTRKQDGTDIITLPCRKIQIINRWQVVGKAMVALPIK 434  
 QY 420 GLIRCSSNITGLLLTRDGGENQTTIEFRPGGDMRDNRWSLYKYKVKIBPLGVAPT 478  
 DB 435 GQIRCSSNITGLLLTRDGGENQTTIEFRPGGDMRDNRWSLYKYKVKIBPLGVAPT 494





HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; M21098; AAA44221.1; -;  
 DR PIR; A31667; VCLJBR.  
 DR PDB; 1IM7; 23-OCT-02.  
 DR HIV; M21098; ENV\$BRVA.  
 DR InterPro; IPR000328; ENV GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;  
 FT SIGNAL 1 30  
 FT CHAIN 31 507  
 FT CHAIN 508 852  
 FT DISULFID 54 74  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 155  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 330  
 FT DISULFID 376 439  
 FT DISULFID 383 412  
 FT CARBOHYD 49 49  
 FT CARBOHYD 88 88  
 FT CARBOHYD 135 135  
 FT CARBOHYD 138 138  
 FT CARBOHYD 154 154  
 FT CARBOHYD 158 158  
 FT CARBOHYD 197 197  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 289 289  
 FT CARBOHYD 295 295  
 FT CARBOHYD 301 301  
 FT CARBOHYD 331 331  
 FT CARBOHYD 354 354  
 FT CARBOHYD 360 360  
 FT CARBOHYD 384 384  
 FT CARBOHYD 390 390  
 FT CARBOHYD 396 396  
 FT CARBOHYD 400 400  
 FT CARBOHYD 442 442  
 FT CARBOHYD 456 456  
 FT CARBOHYD 607 607  
 FT CARBOHYD 612 612  
 FT CARBOHYD 621 621  
 FT CARBOHYD 633 633  
 FT CARBOHYD 670 670  
 FT CARBOHYD 812 812  
 SQ SEQUENCE 852 AA; 2BB866345DEC915F CRC64;  
 Query Match 81.2%; Score 2819.5; DB 1; Length 852;  
 Best Local Similarity 80.8%; Pred. No. 1.3e-215;  
 Matches 529; Conservative 52; Mismatches 43; Indels 31; Gaps 7;  
 3 LWTVYGVVPWKATTTLCFASDAKAYDTAHNVWATHACVPTNPQEVLENVTENF 62  
 34 LWTVYGVVPWKATTTLCFASDAKAYDTAHNVWATHACVPTNPQEVLENVTENF 93  
 63 NMKNMVEQMHEDIISLWDSQSLKPCVKLTPLCVTLNCHDFENATNATNS 122

Db 94 NMKNMVEQMHEDIISLWDSQSLKPCVKLTPLCVTLNCHDFENATNATNS----- 143

Qy 123 RGK----GEVRNCSFNITTSIRDVKVQREYALFYKLDVEPDDNKVNTNTKYLINCNWTS 178

Db 144 -GKMEGEGMKNCNFNITTSIRDVKVQREYALFYKLDVDPDNDK---TNTRYLISCNWTS 199

Qy 179 VITQACPKVGFPEPIHYCTPTGFPALLKCNDDKENGTPCTNVSTVQCTHGIRPVWSTQL 238

Db 200 VITQACPKVGFPEPIHYCAPAGPAIILKCNKKENGTPCTNVSTVQCTHGIRPVWSTQL 259

Qy 239 LLNGSLAEVEVIRSENFNTNAKTIIVQLNVSVSEINCTRNPNHTRKRVTLGPRVWYTTG 298

Db 260 LLNGSLAEVEVIRSENFNTNVKTIIVQLNESVEINCTRNPNHTRKRVTLGPRVWYTTG 319

Qy 299 ILGNIRQAHNCNISRAQNNITLOQIATILRQFNGKTIAFNCSGGDPEIIVHSHFNCGGE 358

Db 320 QIIGDIRAHNCNLSRKWENTLKQIVTKLRVQFNKTIIVFNRSRGGDPEIIVHSHFNCGGE 379

Qy 359 FFYCNSTQLENSAW--NVTSGNTWSVTRKQKDTGDIITLPCRIKQIINRWQVWGKAMVAL 416

Db 380 PFCNTTQLFNSWYRNTGNIT-----EGNSPITLPCRIKQIINRWQVWGKAMVAL 431

Qy 417 PIKGLIRCSNITGLILTRDGGENOT--TEIFRPGGDMRDNWRSLEYKYVVKIEPLG 474

Db 432 PIRQIKCSSNITGLILTRDGGNNETDTEIFRPGGDMRDNWRSLEYKYVVKIEPLG 491

Qy 475 VAPTAKRRVVQREKRAVGLGAMFLGFLGAGSTMGATSMALTVOARQLLSGIVQOQNN 534

Db 492 VAPTAKRRVVQREKRAVG-LGALFLGFLGAGSTMGATSMALTVOARQLLSGIVQOQNN 550

Qy 535 LLRAIKAAQHLQLTVMGIKQIILAVERYLKDQQLLGFQWCGSKLICCTTAVPWNASW 594

Db 551 LLMAIEAQHMLELTVMGIKQIILAVERYLKDQQLLGFQWCGSKLICCTTAVPWNASW 610

Qy 595 SNKTLDOIWNMTWMDREIDNTHLYTLILIESQKQKXQKQELLQDLKWSL 649

Db 611 SNKSLSDIWDNMTWMDREIDNTHLYTLILIESQKQKXQKQELLQDLKWSL 665

RESULT 9

ENV\_HV1SC STANDARD; PRT; 856 AA.

AC P05678;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 15-JUN-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88219542; PubMed=3369091;  
 RA Gurgo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,  
 RA Parrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;  
 RI "Envelope sequences of two new United States HIV-1 isolates."  
 RL Virology 164:531-536(1988).  
 CC -!- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN AIDS PATIENT IN  
 CC 1984 IN SOUTHERN CALIFORNIA.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; M17450; -; NOT ANNOTATED\_CDS.  
 DR HIV; M17450; ENV\$SC.  
 DR InterPro; IPR000328; ENV GP41.



```

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW SIGNAL.
FT CHAIN 1 29
FT CHAIN 30 510
FT CHAIN 511 856
FT SITE 760 760
FT DISULFID 53 73
FT DISULFID 118 206
FT DISULFID 125 197
FT DISULFID 130 160
FT DISULFID 219 247
FT DISULFID 228 239
FT DISULFID 296 330
FT DISULFID 376 439
FT DISULFID 383 412
FT CARBOHYD 87 87
FT CARBOHYD 129 129
FT CARBOHYD 135 135
FT CARBOHYD 140 140
FT CARBOHYD 143 143
FT CARBOHYD 159 159
FT CARBOHYD 163 163
FT CARBOHYD 189 189
FT CARBOHYD 198 198
FT CARBOHYD 234 234
FT CARBOHYD 241 241
FT CARBOHYD 262 262
FT CARBOHYD 276 276
FT CARBOHYD 295 295
FT CARBOHYD 301 301
FT CARBOHYD 302 302
FT CARBOHYD 331 331
FT CARBOHYD 338 338
FT CARBOHYD 354 354
FT CARBOHYD 360 360
FT CARBOHYD 384 384
FT CARBOHYD 394 394
FT CARBOHYD 400 400
FT CARBOHYD 405 405
FT CARBOHYD 442 442
FT CARBOHYD 457 457
FT CARBOHYD 611 611
FT CARBOHYD 616 616
FT CARBOHYD 625 625
FT CARBOHYD 637 637
FT CARBOHYD 674 674
FT CARBOHYD 816 816
SQ SEQUENCE 856 AA; 97055 MW; DAF4DA600BBA7A08 CRC64;

Query Match 81.1%; Score 2818; DB 1; Length 856;
Best Local Similarity 81.3%; Pred. No. 1.8e-215;
Matches 534; Conservative 39; Mismatches 54; Indels 30; Gaps 7;

QY 3 LWTVYVGVVWKEATTTILFCASDAKAYDTBAHWATHACVPTNPDPVLENTENF 62
DB 33 LWTVYVGVVWKEATTTILFCASDAKAYDTBAHWATHACVPTNPDPVLENTENF 92
QY 63 NMWKNMVSQMHEDIISLWDQSLKPCVKLTCLVTLNCTDLNNTNTTSLIIVWVEQ 122
DB 93 NMWKNMVSQMHEDIISLWDQSLKPCVKLTCLVTLNCTDLNNTNTTSLIIVWVEQ 148
QY 123 RGK--GERNCSFNITTSIRDKVOREYALFYKLDVEPIDDKNTNTTKYRLINCNTSV 179
DB 149 RGWEGEGEWNTCSFNITTSIRSKVQKEYALFYKLDVVPID-----NTSYLINCNTSV 201
QY 180 ITQACPKVSFEPIPIHYCTFTGPALLKNDKXKPNGTGCTNVSTVQCTHGIRPVWSTQLL 239
DB 202 ITQACPKVSFEPIPIHYCA-RWFAILCNCKKNGTGPCTNVSTVQCTHGIRPVWSTHLL 260
QY 240 LNSGLAEEVVRSENFTNNAKTIIVQLNVSVEINCTRNPNHTRKVTLPGRVWVTGGE 299

```

---

```

Db 261 LNSGLAEEVVRSENFTNNAKTIIVQLNVSVEINCTRNPNHTRKVTLPGRVWVTGGE 320
QY 300 ILGNIRQAHCNISRAQWNTLQOIAATTIREQNGKTIAPFNQSSGGDPETWHSFNGGGEF 359
Db 321 IIGDIRQAHCNISRAKWNNTLQIVIKLRDQENKTIIFNRSSGGDPETWHSFNGGGEF 380
QY 360 FYCNSTQLFNSAWNT--SNGTWSVTRKQKDTGDIITILPCRIKQIINRQVVGKANYALP 417
Db 381 FYCNSTQLFSSWNGTEGSSNTGG-----NDTITLPCRIKEIINMWQVVGKANYAPP 432
QY 418 IKGLRCSNITGLLLTRDG-----GGENOTTEIFRPGGDMRDNRSELYKYKVKIEP 472
Db 433 IKGQVKCSNITGLLLTRDGGNSKNGSKNENTEIFRPGGDMRDNRSELYKYKVKIEP 492
QY 473 LGVAPTKAKRRVVOREKRAVGMFLGAMFLGAAAGTSMGATSMALTQVQARLLSGIVQOO 532
Db 493 LGVAPTKAKRRVVOREKRAVGTIGAMFLGFLGAAGTSMGATSMALTQVQARLLSGIVQOO 552
QY 533 NNLLRAIKAQOHLLOITVWGIKQLOARILAVERYLKDOQLLGFWCSGKLICTTAVPWA 592
Db 553 NNLLRAIKAQOHLLOITVWGIKQLOARILAVERYLKDOQLLGFWCSGKLICTTAVPWA 612
QY 593 SWSNKTLDQIWNMTWMEWDREIDNTHLYTLIESQNQOEKNOQELQLDKWASL 649
Db 613 SWSNKSLEKINGMTWMEWERIDNTHLYTLIESQNQOEKNOQELQLDKWASL 669

RESULT 10
ENV_HV1A2
ID ENV_HV1A2 STANDARD; PRT; 855 AA.
AC P03378;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (ARV2/SP2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8509453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S., Stempien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A., Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2)".
RL Science 227:484-492(1985).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K02007; AAB59882.1; -.
DR PIR: A03976; VCLJJA2
DR HIV: K02007; ENV3SF2.
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW SIGNAL.
FT CHAIN 1 29
FT CHAIN 30 509
FT CHAIN 510 855
FT DISULFID 53 73
FT DISULFID 118 208

```



FT DISULFID 119 205 BY SIMILARITY.  
 FT DISULFID 126 196 BY SIMILARITY.  
 FT DISULFID 131 157 BY SIMILARITY.  
 FT DISULFID 218 247 BY SIMILARITY.  
 FT DISULFID 228 239 BY SIMILARITY.  
 FT DISULFID 296 331 BY SIMILARITY.  
 FT DISULFID 378 445 BY SIMILARITY.  
 FT DISULFID 385 418 BY SIMILARITY.  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF8107FE0 CRC64;

Query Match  
 Best Local Similarity 82.6%; Pred. No. 8 4e-215; Length 856;  
 Matches 538; Conservative 38; Mismatches 56; Indels 19; Gaps 9;

QY 3 LWVTYVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQVLENTVNF 62  
 DB 34 LWVTYVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQVLENTVNF 93  
 QY 63 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVWEQ 122  
 DB 94 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTSSGRMIME- 150  
 QY 123 RKGEMRNCSEFNITTSIRDKVOREYALFKYLDVEPIDDKNNTNNKRYRLNCNTSVITQ 182  
 DB 151 --KGIIKNCSEFNISIRKVKOYEAFFKLDIIPDN-----DITSYKLNSCNTSVITQ 203  
 QY 183 ACPKVSFPIPIHYCTPTGFALLKNDKFNCTGPTCTNVSTVQCTHGRIPVVSQTLLANG 242  
 DB 204 ACPKVSFPIPIHYCAPAGFALLKCNNTFNCTGPTCTNVSTVQCTHGRIPVVSQTLLANG 263  
 QY 243 SLAESEVIRSENFNTNAKTIIVOLNVSEINCTPRNHRVTL--GPRVWVWTTGEBI 300  
 DB 264 SLAESEVIRSENFNTNAKTIIVOLNVSEINCTPRNHRVTL--GPRVWVWTTGEBI 323  
 QY 301 LGNIQAHCNISRAQWNTLQIATTLRQFG-NKTIAPNQSSGGDPEIVMHSFNCGGEF 359  
 DB 324 -GNMQAHCNISRAQWNTLQIATTLRQFG-NKTIAPNQSSGGDPEIVMHSFNCGGEF 382  
 QY 360 FYCNSQTQFNSAWNTVSGTWSVTRKQDGT-DIITLPCRIKQIINRMQVGVKAMYPALI 418  
 DB 383 FYCNSQTQFNSAWNTVSGTWSVTRKQDGT-DIITLPCRIKQIINRMQVGVKAMYPALI 439

QY 419 KGLIRCSSNITGLLLTRDGGGENTQTEIFRPGGDMRDNRSELKYKVKIEPLGVAFT 478  
 DB 440 SGOIRCSSNITGLLLTRDGGGNNSEIFRPGGDMRDNRSELKYKVKIEPLGVAFT 499  
 QY 479 KAKRRVVOREKAVGMLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOONLLRA 538  
 DB 500 KAKRRVVOREKAVG-IGALFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOONLLRA 558  
 QY 539 IKAQOHLQLTVWGIKQLOARILAVERYLKDQQLLGFWGCSGKLICTTAVPWNASWNTK 598  
 DB 559 IEAQOHLQLTVWGIKQLOARILAVERYLKDQQLLGFWGCSGKLICTTAVPWNASWNTK 618  
 QY 599 LDOIWNMTMWEDEIDNTHLYLTLEESQOQEKQOELLQDKWASL 649  
 DB 619 LDOIWNMTMWEDEIDNTHLYLTLEESQOQEKQOELLQDKWASL 669

RESULT 12  
 ENV\_HV1KB STANDARD; PRT; 861 AA.  
 ID ENV\_HV1KB AC P31819;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-JUL-1993 (Rel. 38, Last annotation update)  
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (KB-1 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=36375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92351552; PubMed=1322587;  
 RA Shimizu H., Hasebe F., Tsuchie H., Morikawa S., Ushijima H., Kitamura T.;  
 RT "Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated transmembrane glycoprotein."  
 RL Virology 189:534-546(1992).  
 CC -!- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR UAG OCCURS BETWEEN THE CODONS FOR 729-ALA AND 730-ARG.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D12582; BAA02124.1; ALT\_SEQ.  
 DR PIR; A42995; VCLJKB.  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.  
 FT SIGNAL 1 35 POTENTIAL.  
 FT CHAIN 36 517 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT CHAIN 518 861 TRANSMEMBRANE GLYCOPROTEIN.  
 FT TRANSMEM 690 711 POTENTIAL.  
 FT DISULFID 59 79 BY SIMILARITY.  
 FT DISULFID 124 212 BY SIMILARITY.  
 FT DISULFID 131 203 BY SIMILARITY.  
 FT DISULFID 136 160 BY SIMILARITY.  
 FT DISULFID 225 254 BY SIMILARITY.  
 FT DISULFID 235 246 BY SIMILARITY.  
 FT DISULFID 303 337 BY SIMILARITY.  
 FT DISULFID 383 446 BY SIMILARITY.  
 FT DISULFID 390 419 BY SIMILARITY.  
 FT CARBOHYD 93 93 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).



FT CARBOHYD 160 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 186 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 197 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 230 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 234 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 241 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 262 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 276 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 289 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 295 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 301 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 332 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 339 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 356 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 386 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 392 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 397 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 406 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 448 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 463 N-LINKED (GLCNAC. . .)  
FT CARBOHYD 611 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 616 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 625 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 637 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 674 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 750 N-LINKED (GLCNAC. . .) (POTENTIAL)  
FT CARBOHYD 816 N-LINKED (GLCNAC. . .) (POTENTIAL)  
SQ SEQUENCE 856 AA; 97224 MW; 0BFBA18931BB27 CRC64;

Query Match 80.8%; Score 2806.5; DB 1; Length 856;  
Best Local Similarity 82.8%; Pred. No. 1.5e-214;  
Matches 539; Conservative 35; Mismatches 58; Indels 19; Gaps 9;

QY 3 LWVTYVYGVVWKEAFTTLFCASDAKAYATEAHNVWATHACVPTNPQEVVLENTENF 62  
DB 34 LWVTYVYGVVWKEAFTTLFCASDAKAYATEAHNVWATHACVPTNPQEVVLENTENF 93

QY 63 NWKKNWVQMEHDIISLDQSLKPCVKLTPLCVTLNCTDNTNTTTELSIIIVWEQ 122  
DB 94 NWKKNWVQMEHDIISLDQSLKPCVKLTPLCVSLKCTDL--KNDTNTSSSGRIME- 150

QY 123 RKGEMRNCSEFNITTSIRDKVOREYALFKLVEPIDDDNKNNTNNTKYRLNCTNSVITQ 182  
DB 151 --KGEIKNCSEFNITTSIRDKVOREYALFKLVEPIDD--DTTSYLTSCNTSVITQ 203

QY 183 ACPKVSFEPIPIHYCTPTGALLKCNKXFGNGTCTNVSTVQCTHGIRPVVSTQLLLNG 242  
DB 204 ACPKVSFEPIPIHYCAPAGFAILLKCNKXFGNGTCTNVSTVQCTHGIRPVVSTQLLLNG 263

QY 243 SLAEEVVRSENFTNAKTIIVQLNVSEINCTRENNTNRK--RVTLPGRVWYTTGEI 300  
DB 264 SLAEEVVRSENFTNAKTIIVQLNVSEINCTRENNTNRKSIRIQRGEGRAFTVIGKI 323

QY 301 LGNIROAHNCISRAOWNNTLQOATTLRBOFG-NKTIAFNQSSGGPPEIWMHSENGGFE 359  
DB 324 -GNMROAHNCISRAOWNNTLQOATTLRBOFG-NKTIAFNQSSGGPPEIWMHSENGGFE 382

QY 360 FYCNSQLFNSAWNTVSTGTSVTRKQKDTG-DIITLPCRIKQIINRWQVGVKAMVALPI 418  
DB 383 FYCNSQLFNSAWNTVSTGTSVTRKQKDTG-DIITLPCRIKQIINRWQVGVKAMVALPI 439

QY 419 KGLIRCSSNITGLLLTRDGGNGQTTEIFRPGGDMRDNRSELYKYKVKVKEPLGVAFT 478  
DB 440 KGLIRCSSNITGLLLTRDGGNGQTTEIFRPGGDMRDNRSELYKYKVKVKEPLGVAFT 499

QY 479 KAKRRVVRQEKRAVGMGLGAFGLGAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538  
DB 500 KAKRRVVRQEKRAVGMGLGAFGLGAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 558

QY 539 IKAQHLLOLTVMGIKQOARILAVERYLKDOQLLGFWSGCKLICTTAVPWNASHSNKT 598  
DB 559 IKAQHLLOLTVMGIKQOARILAVERYLKDOQLLGFWSGCKLICTTAVPWNASHSNKS 618

QY 599 LDOIWNNTWMEWDREIDNYTHLYLIEESQOQKQOELLQDKWASL 649  
DB 619 LEQIWNNTWMEWDREIDNYTHLYLIEESQOQKQOELLQDKWASL 669

RESULT 14  
ENV\_HVLJ3 STANDARD; PRT; 867 AA.

AC PI2489;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DE 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polypeptide GP160 precursor (Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)).  
GN ENV.  
OS Human immunodeficiency virus type 1 (JH3 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11694;  
RN [1] SEQUENCE FROM N.A.  
RP MEDLINE=89352108; PubMed=2669897;  
RX Komyama N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.;  
RT "Nucleotide sequences of gag and env genes of a Japanese isolate of HIV-1 and their expression in bacteria";  
RL AIDS Res. Hum. Retroviruses 5:411-419(1989).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M21138; AAB03526.1; -;  
DR HIV; M21138; ENV5JH3.  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.  
FT SIGNAL 1 30  
FT CHAIN 31 516 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 517 867 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 53 73 BY SIMILARITY.  
FT DISULFID 118 217 BY SIMILARITY.  
FT DISULFID 125 208 BY SIMILARITY.  
FT DISULFID 130 160 BY SIMILARITY.  
FT DISULFID 230 259 BY SIMILARITY.  
FT DISULFID 240 251 BY SIMILARITY.  
FT DISULFID 308 342 BY SIMILARITY.  
FT DISULFID 388 457 BY SIMILARITY.  
FT DISULFID 395 430 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).





*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2004, 09:07:21 ; Search time 21 Seconds  
(without alignments)  
2972.070 Million cell updates/sec

Title: US-09-938-406-1\_COPY\_33\_681  
Perfect score: 3474  
Sequence: 1 ANLWTVYGVVWKEATTT.....QNQKKNQQLLDKWSL 649

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3474	100.0	868	1 VCLJH4	env polyprotein -
2	2869.5	82.6	843	1 H44001	env polyprotein pr
3	2858	82.3	856	1 VCLJ3W	env polyprotein pr
4	2849.5	82.0	847	2 T09448	envelope glycoprot
5	2837	81.7	852	2 T12016	envelope glycoprot
6	2836.5	81.6	847	2 S13289	env protein - huma
7	2816.5	81.1	852	1 VCLJBR	env polyprotein -
8	2815.5	81.0	855	1 VCLJAJ	env polyprotein -
9	2809.5	80.9	854	2 S13288	env protein - huma
10	2808	80.8	856	1 VCLJXX	env polyprotein pr
11	2808	80.8	861	1 VCLJKB	env polyprotein pr
12	2806.5	80.8	856	1 VCLJH3	env polyprotein pr
13	2799.5	80.6	856	1 VCLJVL	env polyprotein pr
14	2795	80.5	861	1 VCLJLV	env polyprotein pr
15	2782	80.1	851	2 S33985	env polyprotein -
16	2765.5	79.6	861	1 VCLJSC	env polyprotein pr
17	2746	79.0	859	1 VCLJMN	env polyprotein pr
18	2623.5	75.5	853	2 S54384	envelope polyprote
19	2616.5	75.3	855	1 VCLJZR	env polyprotein pr
20	2566	73.9	846	1 VCLJND	env polyprotein pr
21	2550	73.4	856	1 A44963	env polyprotein pr
22	2516	72.4	859	2 T01672	envelope polyprote
23	2130.5	61.3	854	1 VCLJSI	env polyprotein pr
24	2074	59.7	506	2 A40218	envelop glycoprote
25	1909	55.0	495	2 S31493	env polyprotein -
26	1825	52.5	443	2 C41621	env polyprotein P
27	1802	51.9	445	2 A41621	env polyprotein M
28	1751.5	50.4	877	2 S49197	envelope protein p
29	1746.5	50.3	454	2 B41621	env polyprotein D

30 1701.5 49.0 863 2 A53034 gag polyprotein -  
31 1204 34.7 297 2 S60538 envelope polyprote  
32 1183.5 34.1 859 1 VCLJST env polyprotein pr  
33 1176 33.9 290 2 S25940 env protein - huma  
34 1168.5 33.6 712 1 VCLJSA env polyprotein pr  
35 1164.5 33.5 858 1 VCLJG2 env polyprotein pr  
36 1157 33.3 852 1 VCLJGG env polyprotein pr  
37 1150.5 33.1 877 2 C46356 env polyprotein -  
38 1140 32.8 859 1 VCLJCT env polyprotein pr  
39 1135.5 32.7 869 2 S33098 envelope polyprote  
40 1134 32.6 881 2 S03068 env protein - huma  
41 1133.5 32.6 864 1 VCLJG4 env polyprotein -  
42 1130 32.5 885 2 S04322 env polyprotein -  
43 1129.5 32.5 294 2 S60545 envelope polyprote  
44 1127.5 32.5 881 1 VCLJG3 env polyprotein -  
45 1127 32.4 301 2 S60548 envelope polyprote

RESULT 1  
VCLJH4  
env polyprotein - human immunodeficiency virus type 1 (isolate CDC-451)  
N/Alternate names: coat polyprotein  
N/Contains: coat protein gp120; coat protein gp41  
C/Species: human immunodeficiency virus type 1, HIV-1  
C/Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 16-Jul-1999  
C/Accession: C25523  
R/Desai, S.M.; Kalyanaraman, V.S.; Casey, J.M.; Srinivasan, A.; Andersen, P.R.; Devare, S  
Proc. Natl. Acad. Sci. U.S.A. 83, 8380-8384, 1986  
A/Title: Molecular cloning and primary nucleotide sequence analysis of a distinct human 3  
A/Reference number: A94136; MUID:87041461; PMID:3490666  
A/Accession: C25523  
A/Molecule type: DNA  
A/Residues: 1-868 <DE>  
A/Cross-references: GB:M13137; NID:G326460; PIDN:AAA44311.1; PID:G326467  
C/Genetics:  
A/Gene: env  
C/Superfamily: type E retrovirus env polyprotein  
C/Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane protei  
F:1-521/Product: coat protein gp120 #status predicted <GP1>  
F:522-868/Product: coat protein gp41 #status predicted <GP2>  
F:89,138,139,142,146,195,198,208,245,252,273,287,300,312,342,349,365,371,395,405,409,459,

ALIGNMENTS

QY 361 YCNSTOLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRIKOIINRWQVGVKAMVALPIKG 420  
 DB 393 YCNSTOLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRIKOIINRWQVGVKAMVALPIKG 452  
 QY 421 LIRCSSNITGLLLTRDGGGNGQTEIFRPGGGDMRDNWRSELYKYKVKIEPLGVAPTGA 480  
 DB 453 LIRCSSNITGLLLTRDGGGNGQTEIFRPGGGDMRDNWRSELYKYKVKIEPLGVAPTGA 512  
 QY 481 KRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTQARQLLSGIVQOQNLLRAIK 540  
 DB 513 KRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTQARQLLSGIVQOQNLLRAIK 572  
 QY 541 AQOHLQLTVMGIKOLOARILAVERYLKDOQLLGFPGWCSGKLICTTAVPWNASNSKTLD 600  
 DB 573 AQOHLQLTVMGIKOLOARILAVERYLKDOQLLGFPGWCSGKLICTTAVPWNASNSKTLD 632  
 QY 601 QIWNMTMMEWDREIDNYTHLIYTLIBESQNOQEKNOQELLQDKWASL 649  
 DB 633 QIWNMTMMEWDREIDNYTHLIYTLIBESQNOQEKNOQELLQDKWASL 681  
 RESULT 2  
 H44001  
 env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)  
 N;Alternate names: coat polyprotein  
 N;Contains: coat protein gp120; coat protein gp41  
 C;Species: human immunodeficiency virus type 1, HIV-1  
 A;Note: host Homo sapiens (man)  
 C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Sep-1994  
 C;Accession: H44001  
 J;Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.  
 J. Virol. 66, 6587-6600, 1992  
 A;Title: Complete nucleotide sequence, genome organization, and biological properties of  
 A;Reference number: A44001; MUID:93021387; PMID:1404605  
 A;Accession: H44001  
 A;Molecule type: DNA  
 A;Residues: 1-843 <LTY>  
 A;Cross-references: GB:M32358  
 C;Genetics:  
 A;Gene: env  
 C;Superfamily: type E retrovirus env polyprotein  
 C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
 F;1-29/Domain: signal sequence #status predicted <SIG>  
 F;19-35/Region: hydrophobic  
 F;30-489/Product: coat protein gp120 #status predicted <GP1>  
 F;490-843/Product: coat protein gp41 #status predicted <GP2>  
 F;499-515/Region: hydrophobic  
 F;673-689/Region: hydrophobic  
 F;738-755/Domain: transmembrane #status predicted <TMN>  
 F;87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

Query Match 82.6%; Score 2869.5; DB 1; Length 843;  
 Best Local Similarity 83.5%; Pred. No. 5.7e-203;  
 Matches 541; Conservative 35; Mismatches 47; Indels 25; Gaps 6;  
 QY 3 LWTVYGVVPWKATTTLCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 62  
 DB 33 LWTVYGVVPWKATTTLCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 92  
 QY 63 NMWKNNVQEMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNTNTTSLSIIVVWEQ 122  
 DB 93 NMWKNNVQEMHEDIISLWQSLKPCVKLTPLCVTLNCTDL--RNAINTSSS----WET 146  
 QY 123 RGKEMRNCSPNITTSIRDKVQREYALFYKLDVEPIDDNKNTNTKYRLINCMNTSVITQ 182  
 DB 147 MEKGEIKNCSPNITTSIRDKVQREYALFYKLDVEPID--NASTYRLINCMNTSVITQ 199  
 QY 183 ACQKVSFEPIPIHYCTPTGFALLKCNCKNGTGPCNTNVSTVQCTHGIRPVVSTQILLG 242  
 DB 200 ACQKVSFEPIPIHYCAPAGFALLKCNCKNGTGPCNTNVSTVQCTHGIRPVVSTQILLG 259  
 QY 243 SLAEERVIRSENFNTNNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGFGRWYTTGILG 302

DB 260 SLAEERVIRSENFNTNNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGFGRWYTTGILG 319  
 QY 303 NIRAHCNISRQWNTLQOIATLREOFG-NKTIARNSQSSGDPRIVMSFNGCGGEFF 361  
 DB 320 DIRQAHCNLSKQWNTLEQIAIKLKEQFGNNKTIIFNPSSGGDPRIVMSFNGCGGEFF 379  
 QY 362 CNSTOLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRIKOIINRWQVGVKAMVALPIKGL 421  
 DB 380 CNSTOLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRIKOIINRWQVGVKAMVALPIKGL 429  
 QY 422 IRCSSNITGLLLTRDGGGNGQTEIFRPGGGDMRDNWRSELYKYKVKIEPLGVAPTGA 481  
 DB 430 IRCSSNITGLLLTRDGGGNGQTEIFRPGGGDMRDNWRSELYKYKVKIEPLGVAPTGA 489  
 QY 482 KRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTQARQLLSGIVQOQNLLRAIK 541  
 DB 490 KRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTQARQLLSGIVQOQNLLRAIK 548  
 QY 542 AQOHLQLTVMGIKOLOARILAVERYLKDOQLLGFPGWCSGKLICTTAVPWNASNSKTLD 601  
 DB 549 AQOHLQLTVMGIKOLOARILAVERYLKDOQLLGFPGWCSGKLICTTAVPWNASNSKTLD 608  
 QY 602 IWNMTMMEWDREIDNYTHLIYTLIBESQNOQEKNOQELLQDKWASL 649  
 DB 609 IWNMTMMEWDREIDNYTHLIYTLIBESQNOQEKNOQELLQDKWASL 656  
 RESULT 3  
 VCLJ3W  
 env polyprotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)  
 N;Contains: coat protein gp120; coat protein gp41  
 C;Species: human immunodeficiency virus type 1, HIV-1  
 C;Date: 30-Sep-1988 #sequence\_revision 30-Sep-1988 #text\_change 07-Nov-1997  
 C;Accession: A24774  
 R;Stargich, B.R.; Hahn, B.H.; Shaw, G.M.; McNeely, P.D.; Modrow, S.; Wolf, H.; Parks, E.S.  
 Cell 45, 637-648, 1986  
 A;Title: Identification and characterization of conserved and variable regions in the en  
 A;Reference number: A24774; MUID:86218077; PMID:2423250  
 A;Accession: A24774  
 A;Molecule type: DNA  
 A;Residues: 1-856 <STA>  
 A;Cross-references: GB:K03455; GB:M38432; NID:g1906382  
 C;Genetics:  
 A;Gene: env  
 C;Superfamily: type E retrovirus env polyprotein  
 C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
 F;1-29/Domain: signal sequence #status predicted <Sig>  
 F;30-501/Product: coat protein gp120 #status predicted <GP1>  
 F;502-847/Product: coat protein gp41 #status predicted <GP2>  
 F;87,134,140,151,155,183,197,234,241,262,276,289,295,331,338,354,360,390,394,404,447,459,

Query Match 82.3%; Score 2858; DB 1; Length 856;  
 Best Local Similarity 82.8%; Pred. No. 4.1e-202;  
 Matches 539; Conservative 34; Mismatches 60; Indels 18; Gaps 5;  
 QY 3 LWTVYGVVPWKATTTLCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 62  
 DB 33 LWTVYGVVPWKATTTLCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 92  
 QY 63 NMWKNNVQEMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNTNTTSLSIIVVWEQ 122  
 DB 93 NMWKNNVQEMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNTNTTSLSIIVVWEQ 140  
 QY 123 R---GKEMRNCSPNITTSIRDKVQREYALFYKLDVEPIDDNKNTNTKYRLINCMNTSV 179  
 DB 141 KTIIGGEVKNCSFNITTSIRDKVQREYALFYKLDVEPIDDNKNTNTKYRLINCMNTSV 200  
 QY 180 ITOACQKVSFEPIPIHYCTPTGFALLKCNCKNGTGPCNTNVSTVQCTHGIRPVVSTQILL 239  
 DB 201 ITOACQKVSFEPIPIHYCAPAGFALLKCNCKNGTGPCNTNVSTVQCTHGIRPVVSTQILL 260  
 QY 240 LNSLAEEFVIRSENFNTNNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGFGRWYTTG 298

Db 261 LNSLAEEVIRSENFTDNAKTIIIVHLNESVEINCTRNNNVRRBHHIIGPGAFY-TG 319  
QY 299 EILGNTRQAHCHNISRAQWNTTQQIATTIREQFNGKTIAPFNSSGGDPEIVMHSFNCGE 358  
Db 320 EIRGNTRQAHCHNISRAQWNTTQQIATTIREQFNGKTIAPFNSSGGDPEIVMHSFNCGE 379  
QY 359 FFYCNSTQFNQWNTTSGTWSVTRKQDGTDIITLPCRIKQIINRWQVGVKAMVALPI 418  
Db 380 FYCDSTQFNQWNTTSGTWSVTRKQDGTDIITLPCRIKQIINRWQVGVKAMVALPI 438  
QY 419 KGLIRCSSNITGLLLTRDGGGNGQTTTFRPGGDMRDNRWSELYKYKVVKIEPLGVAPT 478  
Db 439 GGOIRCSSNITGLLLTRDGGGNGQTTTFRPGGDMRDNRWSELYKYKVVKIEPLGVAPT 498  
QY 479 KAKRRVVQREKRAVGMGLGFLGAAGSTMGATSMALTQVQARLLSGIVQOQNLLRA 538  
Db 499 KAKRRVVQREKRAVGMGLGFLGAAGSTMGATSMALTQVQARLLSGIVQOQNLLRA 558  
QY 539 IKAQCHLLQLTVMGIKQIOLARILAVERYLKDQQLGFWGCSGKLICTTAVPWNASWNSKT 598  
Db 559 IEAQCHLLQLTVMGIKQIOLARILAVERYLKDQQLGFWGCSGKLICTTAVPWNASWNSKT 618  
QY 599 LDQIWNMTWMEWRIDNVTYLIYTLIBESQOQEKNOQELLQDKWASL 649  
Db 619 MQQIWNMTWMEWRIDNVTYLIYTLIBESQOQEKNOQELLQDKWASL 669

RESULT 4  
T09448  
envelope glycoprotein - human immunodeficiency virus type 1 (strain JRFL)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 24-Nov-1999  
C:Accession: T09448  
Rifang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,  
submitted to the EMBL Data Library, July 1996  
A:Reference number: Z16673  
A:Accession: T09448  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-847 <PAN>  
C:Cross-references: EMBL:U63632; NID:g1465777; PID:g1465781  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein

Query Match 82.0%; Score 2849.5; DB 2; Length 847;  
Best Local Similarity 82.8%; Pred. No. 1.7e-201;  
Matches 536; Conservative 40; Mismatches 52; Indels 19; Gaps 4;  
QY 3 LMTVYVYGVPMKAEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 62  
Db 33 LMTVYVYGVPMKAEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 92  
QY 63 NNMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTNTTSLIIVWBEQ 122  
Db 93 NNMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTNTTSLIIVWBEQ 146  
QY 123 RKGEMRNCSEFNITTSIRKQVREYALFYKLDVEPIDDKNNTNTNTKYRLINCNTSVITQ 182  
Db 147 -ERGBIKNCSEFNITTSIRKQVREYALFYKLDVEPIDDKNNTNTNTKYRLINCNTSVITQ 200  
QY 183 ACQKVSFEPIHYCTPTGPFALLKNDKFKNGTGCTNVSTVQCTHGIRPVVSTQLLNG 242  
Db 201 ACQKVSFEPIHYCTPTGPFALLKNDKFKNGTGCTNVSTVQCTHGIRPVVSTQLLNG 260  
QY 243 SLAEEVIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGCRVWYTTG 302  
Db 261 SLAEEVIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGCRVWYTTG 320  
QY 303 NTRQAHCHNISRAQWNTTQQIATTIREQFNGKTIAPFNSSGGDPEIVMHSFNCGEFFYC 362  
Db 321 DIRQAHCHNISRAQWNTTQQIATTIREQFNGKTIAPFNSSGGDPEIVMHSFNCGEFFYC 380

QY 363 NSTQFNQWNTTSGTWSVTRKQDGTDIITLPCRIKQIINRWQVGVKAMVALPIKGLI 422  
Db 381 NSTQFNQWNTTSGTWSVTRKQDGTDIITLPCRIKQIINRWQVGVKAMVALPIKGLI 434  
QY 423 RCSSNITGLLLTRDGGGNGQTTTFRPGGDMRDNRWSELYKYKVVKIEPLGVAPTAKR 482  
Db 435 RCSSNITGLLLTRDGGGNGQTTTFRPGGDMRDNRWSELYKYKVVKIEPLGVAPTAKR 494  
QY 483 RVQREKRAVGMGLGFLGAAGSTMGATSMALTQVQARLLSGIVQOQNLLRAKQ 542  
Db 495 RVQREKRAVGMGLGFLGAAGSTMGATSMALTQVQARLLSGIVQOQNLLRAKQ 553  
QY 543 QHLLQLTVMGIKQIOLARILAVERYLKDQQLGFWGCSGKLICTTAVPWNASWNSKTLDQI 602  
Db 554 QHLLQLTVMGIKQIOLARILAVERYLKDQQLGFWGCSGKLICTTAVPWNASWNSKTLDQI 613  
QY 603 WNMNTWMEWRIDNVTYLIYTLIBESQOQEKNOQELLQDKWASL 649  
Db 614 WNMNTWMEWRIDNVTYLIYTLIBESQOQEKNOQELLQDKWASL 660

RESULT 5  
T12016  
envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T12016  
R.McCutchan, F.B.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.  
AIDS Res. Hum. Retroviruses 14, 329-337, 1998  
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S  
A:Reference number: Z17379; MUID:9817816; PMID:9519894  
A:Accession: T12016  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-852 <MCC>  
C:Cross-references: EMBL:U09034; NID:g2351783; PID:AA59271.1; PID:g2351784  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein

Query Match 81.7%; Score 2837; DB 2; Length 852;  
Best Local Similarity 83.1%; Pred. No. 1.4e-200;  
Matches 541; Conservative 35; Mismatches 53; Indels 22; Gaps 8;  
QY 3 LMTVYVYGVPMKAEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 62  
Db 33 LMTVYVYGVPMKAEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 92  
QY 63 NNMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTNTTSLIIVWBEQ 122  
Db 93 NNMKNMVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTNTTSLIIVWBEQ 143  
QY 123 RG---KGMRCNCSFNITTSIRKQVREYALFYKLDVEPIDDKNNTNTNTKYRLINCNTS 178  
Db 144 NKGMEGEIKNCSEFNITTSIRKQVREYALFYKLDVEPIDDKNNTNTNTKYRLINCNTS 198  
QY 179 VITQACPKVSFEPIHYCTPTGPFALLKNDKFKNGTGCTNVSTVQCTHGIRPVVSTQ 238  
Db 199 VITQACPKVSFEPIHYCTPTGPFALLKNDKFKNGTGCTNVSTVQCTHGIRPVVSTQ 258  
QY 239 LINGSLAEEVIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGCRVWYTTG 298  
Db 259 LINGSLAEEVIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGCRVWYTTG 318  
QY 299 ETLGNIRQAHCHNISRAQWNTTQQIATTIREQFNGKTIAPFNSSGGDPEIVMHSFNCGE 358  
Db 319 ALIGDIRQAHCHNISRAQWNTTQQIATTIREQFNGKTIAPFNSSGGDPEIVMHSFNCGE 378  
QY 359 FFYCNSTQFNQWNTTSGTWSVTRKQDGTDIITLPCRIKQIINRWQVGVKAMVALPI 418  
Db 379 FFYCNSTQFNQWNTTSGTWSVTRKQDGTDIITLPCRIKQIINRWQVGVKAMVALPI 435  
QY 419 KGLIRCSSNITGLLLTRDGGGNGQTTTFRPGGDMRDNRWSELYKYKVVKIEPLGVAPT 478

Db 436 RGQIRCSNITGILLTRDGGTNNSTNETFRPGGDMRDNWRSELYKYKVKIEPLGVAPT 495  
Qy 479 KAKRRVVQREKRAVGMFLGAMFLGAGSTMGATSMALTVQARQLLSGIVQQNNLLRA 538  
Db 496 KAKRRVVQREKRAVG-IGALFLGFLGAGSTMGAAVTLTVQARQLLSGIVQQNNLLRA 554  
Qy 539 IKAQOHLQLTVMGIKOLQARILAVERYLKDQQLLGFWSGSKLICCTTAVPWNASWKNKT 598  
Db 555 IEAQOHLQLTVMGIKOLQARILAVERYLKDQQLLGFWSGSKLICCTTAVPWNASWKNKS 614  
Qy 599 LDQIWNNTWMEWDREINDYTHLYTILIESQNEQKNOQELLQDLKWSL 649  
Db 615 LDKIWNNTWMEWDREINDYTHLYTILIESQNEQKNOQELLQDLKWSL 665

RESULT 6  
S13289  
env protein - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
C:Accession: S13289  
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.  
Nature 348, 69-73, 1990  
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120  
A:Reference number: S13288; MUID:91043044; PMID:2172833  
A:Accession: S13289  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-847 <ORF>  
C:Superfamily: type E retrovirus env polyprotein

Query Match 81.6%; Score 2836.5; DB 2; Length 847;  
Best Local Similarity 82.5%; Pred. No. 1.5e-200;  
Matches 534; Conservative 39; Mismatches 55; Indels 19; Gaps 4;

Qy 3 LWTVVYGVVPVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 62  
Db 33 LWTVVYGVVPVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 92  
Qy 63 NMWKNWVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTNTLSIIVWVEQ 122  
Db 93 NMWKNWVEQMHEDIISLDQSLKPCVKLTPLCVTLNCKDVNATNTDSEGTM----- 146

Qy 123 RGKEMRNCSEFNITTSIRDKVQREYALFYKLDVPEIDDNKNNTNTKYLINCMNTS 182  
Db 147 -ERGEIKNCSEFNITTSIRDEYQKEVALFYKLDVPEIDN-----NNTSYRLISCDSVTITQ 200

Qy 183 ACPKVSPEPIPIHYCTPTGFPALLKCNCKKFGTGPCTNVSTVQCTHGIRPVVSTQLLNG 242  
Db 201 ACPKISPEPIPIHYCAPAGFAILLKCNCKTFTNGKGPCKNVSTVQCTHGIRPVVSTQLLNG 260

Qy 243 SLAESEVVRSENFNNAKTIIIVQLNVSEINCRPNNHTRKRVTLGPRVYMTTGEILG 302  
Db 261 SLAESEVVRSENFNNAKTIIIVQLKESVEINCRPNNTRKSHIIGPGRAYTITGEIIG 320

Qy 303 NIRAHCNISRAQNNLTQIATTLREDFGNKTTAFNOSGGDPEIIVMHSFNCGEFFYC 362  
Db 321 DIRQAHCNISRAKNDTLKQIVIKLREDFGNKTTAFNHSFGDPEIIVMHSFNCGEFFYC 380

Qy 363 NSTQLFSAWNTSGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGMKAMVALPIKGLI 422  
Db 381 NSTQLNNTSWNTEGS-----NNTGNTITLPCRIKQFINWQEVGMKAMVPIRQOI 434

Qy 423 RCSSNITGLLTRDGGGNNOTTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPTAKR 482  
Db 435 RCSSNITGLLTRDGGINGENTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPTAKR 494

Qy 483 RVQREKRAVGMFLGFLGAGSTMGATSMALTVQARQLLSGIVQQNNLLRAIKAO 542  
Db 495 RVQREKRAVG-IGAVFLGFLGAGSTMGAAVTLTVQARQLLSGIVQQNNLLRAIEAQ 553

Qy 543 QHLLQLTVMGIKOLQARILAVERYLKDQQLLGFWSGSKLICCTTAVPWNASWKNKTLDOI 602

Db 554 QRMQLTVMGIKOLQARILAVERYLGDQQLLGFWSGSKLICCTTAVPWNASWKNKSLDRI 613  
Qy 603 WNNMTWMEWDREINDYTHLYTILIESQNEQKNOQELLQDLKWSL 649  
Db 614 WNNMTWMEWDREINDYTHLYTILIESQNEQKNOQELLQDLKWSL 660

RESULT 7  
VCLJBR  
env polyprotein - human immunodeficiency virus type 1 (isolate BR)  
N:Alternate names: coat polyprotein  
N:Contains: coat protein gp120; coat protein gp41  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 16-Feb-1997  
C:Accession: A31667  
R:Anand, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar, S.  
Virology 168, 79-89, 1989  
A:Title: Biological and molecular characterization of human immunodeficiency virus (HIV-1)  
A:Reference number: A94389; MUID:89085613; PMID:2789516  
A:Accession: A31667  
A:Molecule type: DNA  
A:Residues: 1-852 <ANA>  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: capsid protein; coat protein; polyprotein; transmembrane protein  
F:1-516/Product: coat protein gp120 #status predicted <CPI>  
F:517-852/Product: coat protein gp41 #status predicted <CF2>

Query Match 81.1%; Score 2816.5; DB 1; Length 852;  
Best Local Similarity 80.6%; Pred. No. 4.6e-199;  
Matches 528; Conservative 53; Mismatches 43; Indels 31; Gaps 7;

Qy 3 LWTVVYGVVPVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 62  
Db 34 LWTVVYGVVPVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 93

Qy 63 NMWKNWVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTNTLSIIVWVEQ 122  
Db 94 NMWKNWVEQMHEDIISLDQSLKPCVKLTPLCVTLNCHDFNATNATNS----- 143

Qy 123 RGK-----GEVRNCSEFNITTSIRDKVQREYALFYKLDVPEIDDNKNNTNTKYLINCMNTS 178  
Db 144 -GKWMGEGEMKNCSEFNITTSIRDKWQKEYALFYKLDVPEIDNK---TNTYRLISCNTS 199

Qy 179 VITQACPKVSPEPIPIHYCTPTGFPALLKCNCKKFGTGPCTNVSTVQCTHGIRPVVSTQL 238  
Db 200 VITQACPKVTFPEPIPIHYCAPAGFAILLKCNCKKFGTGPCTNVSTVQCTHGIRPVVSTQL 259

Qy 239 LINGSLAESEVVRSENFNNAKTIIIVQLNVSEINCRPNNHTRKRVTLGPRVYMTTGT 298  
Db 260 LINGSLAESEVVRSENFNNAKTIIIVQLNESVEINCRPNNTRKRVTLGPRVYMTTGT 319

Qy 299 EILGNIRQAHCNISRAQNNLTQIATTLREDFGNKTTAFNOSGGDPEIIVMHSFNCGE 358  
Db 320 QIIGDIRRAHCNLSRKWENTLKQIVTKLRVQPKNTKIVFNHSSGGDPEIIVMHSFNCGE 379

Qy 359 PFYCNSTQLFNSAW-NVTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGMKAMVAL 416  
Db 380 PFCNTTQLFNSTYRNTTGNIT-----EGNSPITLPCRIKQIINRWQEVGMKAMVAP 431

Qy 417 PIKGLIRCSNITGLLTRDGGGNNOT--TEIFRPGGDMRDNWRSELYKYKVKIEPLG 474  
Db 432 PIRQIKCSSNITGLLTRDGGNNETDTEIFRPGGDMRDNWRSELYKYKVKIEPLG 491

Qy 475 VAPTAKRRVVQREKRAVGMFLGFLGAGSTMGATSMALTVQARQLLSGIVQQNN 534  
Db 492 VAPTAKRRVVQREKRAVG-LGALFLGFLGAGSTMGAAVTLTVQARQLLSGIVQQNN 550

Qy 535 LLRAIKAOQHLLQLTVMGIKOLQARILAVERYLKDQQLLGFWSGSKLICCTTAVPWNASW 594  
Db 551 LLMAIEAQOQHMLLETVNGIKOLQARILAVERYLKDQQLLGFWSGSKLICCTTAVPWNASW 610

QY 595 SNKTLDOINWNTWMEWDREIDNYTHLYIYTLIEESQNOEKQOELLQDKWASL 649  
Db 611 SNKSLSDIWNNTWMEWDREIDNYTHLYIYTLIEESQIOEKNEKELLELDKWASL 665

RESULT 8  
VCLJ2A  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate ARV-2)  
N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03976  
R:Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown-SH  
Science 227, 484-492, 1985  
A:Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).  
A:Reference number: A04003; MUID:85090453; PMID:2578227  
A:Accession: A03976  
A:Molecule type: DNA  
A:Residues: 1-855 <SAN>  
A:Cross-references: GB:K02007; NID:G328658; PIDN:AAB59882.1; PID:G328666  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro  
F:1-30/Domain: signal sequence #status predicted <Sig>  
F:31-509/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:510-855/Product: transmembrane glycoprotein #status predicted <TM>  
F:87,129,140,158,184,190,200,244,265,292,298,304,334,341,359,364,388,394,400,408,445,458  
F:610,624,636,815/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 81.0%; Score 2815.5; DB 1; Length 855;  
Best Local Similarity 81.2%; Pred. No. 5.4e-199;  
Matches 528; Conservative 47; Mismatches 58; Indels 17; Gaps 6;

QY 3 LWVTYVYGVPMKEATTTLCFASDAKAYDTEAHNVWATHACVPTNPNQEVVLNVTFN 62  
Db 33 LWVTYVYGVPMKEATTTLCFASDAKAYDTEAHNVWATHACVPTDNPQEVVLNVTFN 92

QY 63 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVWEQ 122  
Db 93 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTSSN-----WKE 146

QY 123 RKGEMRNCSPNITTSIRDKVOREYALFYKLDVEPIDDKNTNNTKYRLINCNTSVITQ 182  
Db 147 EIKGEIKNCSPNITTSIRDKIQENALFNLVDVPIDNATTTNTYNYRLIHCNRSVITQ 206

QY 183 ACPKVSFEPIPIHYCTPTGTFALLKCNKKFNGTGPCTNVSTVQCTHGIRPVYSTQLLNG 242  
Db 207 ACPKVSFEPIPIHYCTPTGTFALLKCNKKFNGTGPCTNVSTVQCTHGIRPIYSTQLLNG 266

QY 243 SLAEVWIRSENFNTNAKTIIVQLNVSVIEINCTRPNNHTRKRVTLGPRVWYTTGEILG 302  
Db 267 SLAEVWIRSENFNTNAKTIIVQLNVSVIEINCTRPNNHTRKSIYIGPGRAFTTGTIG 326

QY 303 NTRQAHCHNSRAQWNTLQOIATTLRBOFG-NKTIAPNOSGGDPPEIVMHSFNCGEPPY 361  
Db 327 DIRXAHCHNSRAQWNTLEQIVKLRQEPGNKTIIVFNOSGGDPPEIVMHSFNCGEPPY 386

QY 362 CNSTOLFNSAWNT-SNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKAWYALPIKG 420  
Db 387 CNTQTFNNTWRLNHTEGT-----KNDT---IILPCRIKQIINRWQVVGKAWYALPIGG 438

QY 421 LIRCSSNITGLLITRDGG-GENQTTBIFRPGGDMRDNRSSELYKYKVKIPELGVAPTK 479  
Db 439 QISCSSNITGLLITRDGGVINDTEVFRPGGDMRDNRSSELYKYKVKIPELGVAPTK 498

QY 480 AKRRVQREKRAVGMGLGAMFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOQNLLRAI 539  
Db 499 AKRRVQREKRAVGIVGAMFLGFLGAAGSTMGVSLLTVOARQLLSGIVQOQNLLRAI 558

QY 540 KAOQHLLQLTWGIKQOLQARILAVERYLKDQQLLGFWCGSKLICCTTAVPNWASNSKSL 599  
Db 540 KAOQHLLQLTWGIKQOLQARILAVERYLKDQQLLGFWCGSKLICCTTAVPNWASNSKSL 599

Db 559 EAQOHLQLTWGIKQOLQARVLAVERYLKDQQLLGFWCGSKLICCTTAVPNWASNSKSL 618  
QY 600 DOIWNNTWMEWDREIDNYTHLYIYTLIEESQNOEKQOELLQDKWASL 649  
Db 619 EDIWNNTWMEWDREIDNYTHLYIYTLIEESQNOEKQOELLQDKWASL 668

RESULT 9  
SI3288  
env protein - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
C:Accession: SI3288  
R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Digne, A.; Idler, K.; Zack, J.A.  
Nature 348, 69-73, 1990  
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120 c  
A:Reference number: SI3288; MUID:91043044; PMID:2172833  
A:Accession: SI3288  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-854 <OBR>  
C:Superfamily: type E retrovirus env polyprotein

Query Match 80.9%; Score 2809.5; DB 2; Length 854;  
Best Local Similarity 82.8%; Pred. No. 1.5e-198;  
Matches 539; Conservative 34; Mismatches 57; Indels 21; Gaps 9;

QY 3 LWVTYVYGVPMKEATTTLCFASDAKAYDTEAHNVWATHACVPTNPNQEVVLNVTFN 62  
Db 34 LWVTYVYGVPMKEATTTLCFASDAKAYDTEAHNVWATHACVPTDNPQEVVLNVTFN 93

QY 63 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVWEQ 122  
Db 94 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDL--KNDTNTNSSSGRMIME- 150

QY 123 RKGEMRNCSPNITTSIRDKVOREYALFYKLDVEPIDDKNTNNTKYRLINCNTSVITQ 182  
Db 151 --KGEIKNCSPNITTSIRDKVOREYALFYKLDVEPID--NTSYRLINCNTSVITQ 201

QY 183 ACPKVSFEPIPIHYCTPTGTFALLKCNKKFNGTGPCTNVSTVQCTHGIRPVYSTQLLNG 242  
Db 202 ACPKVSFEPIPIHYCAPAGFALLKCNKNFTNGTGPCTNVSTVQCTHGIRPVYSTQLLNG 261

QY 243 SLAEVWIRSENFNTNAKTIIVQLNVSVIEINCTRPNNHTRK--RVTLGPRVWYTTGEI 300  
Db 262 SLAEVWIRSENFNTNAKTIIVQLNVSVIEINCTRPNNHTRKSIYIGPGRAFTTGTIG 321

QY 301 LGNIRQAHCHNSRAQWNTLQOIATTLRBOFG-NKTIAPNOSGGDPPEIVMHSFNCGEFF 359  
Db 322 -GNMROAHCHNSRAQWNTLQOIATTLRBOFG-NKTIAPNOSGGDPPEIVMHSFNCGEFF 380

QY 360 FYCNSTOLFNSAWNTSNGTWSVTRKQKDTG-DIITLPCRIKQIINRWQVVGKAWYALPI 418  
Db 381 FYCNSTOLFNSWTM---FNSTWSTEGSNNTGSDTTITLPCRIKQFINMWQVVGKAWYAPPI 437

QY 419 KGLIRCSSNITGLLITRDGGGNQTTBIFRPGGDMRDNRSSELYKYKVKIPELGVAPTK 478  
Db 438 SQGIRCSSNITGLLITRDGGNNNGSEIFRPGGDMRDNRSSELYKYKVKIPELGVAPTK 497

QY 479 KAKRRVQREKRAVGMGLGAMFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOQNLLRA 538  
Db 498 KAKRRVQREKRAVG-IGALFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOQNLLRA 556

QY 539 IKAQOHLQLTWGIKQOLQARILAVERYLKDQQLLGFWCGSKLICCTTAVPNWASNSKNT 598  
Db 557 IEAQOHLQLTWGIKQOLQARILAVERYLKDQQLLGFWCGSKLICCTTAVPNWASNSKS 616

QY 599 LDOIWNNTWMEWDREIDNYTHLYIYTLIEESQNOEKQOELLQDKWASL 649  
Db 617 LDOIWNNTWMEWDREIDNYTHLYIYTLIEESQNOEKQOELLQDKWASL 667

RESULT 10

## VCLJKB

env polyprotein precursor - human immunodeficiency virus type 1 (strain KB-1-gp32)  
N:Alternate names: coat polyprotein  
N:Contains: coat protein gp120; coat protein gp32  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 01-Mar-1996  
C:Accession: B42995  
R:Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura, T.  
Virology 189, 534-546, 1992  
A:Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated  
A:Reference number: A42995; MUID:92351552; PMID:1322587  
A:Accession: B42995  
A:Molecule type: mRNA  
A:Residues: 1-729 <SHI>  
A:Cross-references: GB:S41266; GB:D01206  
C:Genetics:  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
F:1-689/Domain: extracellular #status predicted <EXT>  
F:1-33/Domain: signal sequence #status predicted <SIG>  
F:17-33/Region: hydrophobic #status predicted  
F:34-517/Product: coat protein gp120 #status predicted <CP1>  
F:514-517/Region: cleavage processing #status predicted  
F:518-729/Product: coat protein gp32 #status predicted <CP2>  
F:518-534/Region: hydrophobic #status predicted  
F:690-711/Domain: transmembrane #status predicted <TM1>  
F:712-729/Domain: intracellular #status predicted <INT>  
F:793,141,145,146,163,191,192,237,241,248,269,283,296,308,338,345,361,367,397,403,408,414

Query Match 80.8%; Score 2808; DB 1; Length 729;  
Best Local Similarity 81.8%; Pred. No. 1.6e-198;  
Matches 534; Conservative 49; Mismatches 48; Indels 22; Gaps 9;  
Qy 3 LWTVYGVVPWKATTLFCASDAKAYDTEAHNVWATHACVPDPNPQEVVLNVTFN 62  
Db 39 LWTVYGVVPWKATTLFCASDAKAYDTEAHNVWATHACVPDPNPQEVVLNVTFN 98  
Qy 63 NMWKNVVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNTNTNTTSLIIVVWEQ 122  
Db 99 NMWKNVVEQMHENIISLDQSLKPCVKLTPLCVTLNCTDL--RNTTNNS-SI---EE 151  
Qy 123 RGKGMNCSFNITTSIRDKVQREYALFYKLDVPEIDDNKNNTNTKYRLNCSVTIQ 182  
Db 152 KMKGEIKNCSFNVTNIRDKVQREYALFYKLDVPEIDDNKNST-NTCYRLISCDTSVITQ 210  
Qy 183 ACPKVSPEPIPIHYCTPTGTFALLKCNCKKFGTGPCTNVSTVQCTHGRIPVYSTQLLNG 242  
Db 211 ACPKVSPEPIPIHYCTPAGFALLKCNCKNTFNGTGPCKNVSTVQCTHGRIPVYSTQLLNG 270  
Qy 243 SLAEEVVRSENFNTNNAKTIIVQLNVSVETNCPNHNTRKRVTLGPRVWYTTGILG 302  
Db 271 SLAEEVVRSENFNTDNVKTIIVQLNETVKINCIRPNKTKRVTMGPRVYTTGILG 330  
Qy 303 NROAHCHNISRAQWNTLQOIATTLREQFGNKTIAFNQSOGDDEIIVMHSFNCGGEFFYC 362  
Db 331 DIRAHCHNISRAEWNTLEQIANLKRQFENKTIIVFNQSOGDDEIIVMHSFNCGGEFFYC 390  
Qy 363 NSTOLFNSAMNVTNGTWSVTRKQDGT-DIITLPCRIKQIINWQVGVKAMYPALPIKGL 421  
Db 391 DSSQLENSTH--LSNGTW-----WNGTGPENITLPCRIKQIINWQVGVKAMYPALPIKGL 443  
Qy 422 IRCSSNTIGLLTRDGG-----GENOTTEIFRPGGDMRDNWRSELYKYKVKIEPLGVA 476  
Db 444 IRCSSNTIGLLTRDGGNTNNSSIEIFRPGGDMRDNWRSELYKYKVKIEPLGVA 503  
Qy 477 PTKAKRRVQREKAVGMLGAMFLGFLAAGSTWGAATSMALTVOARQLLSGIVQOQNLL 536  
Db 504 FTRAKRRVQREKAVG--IGAVFLGFLAAGSTWGAATVLTVOARQLLPVIVQOQNLL 562  
Qy 537 RAIKAQOHLIQLTVWGIKQOLARILAVERYIKDQQLIGFVGCSCGLICTTAVPWNASWN 596

Db 563 RAIDAQOHLIQLTVWGIKQOLARILAVERYIKDQQLIGFVGCSCGLICTTAVPWNASWN 622  
Qy 597 KTLDOINNNMTMWEDEIDNYTHLIYTLIEESQOQEKNOQELLOLDKWSL 649  
Db 623 KSFNEIWDNMTMWEDEINNYTHLIYTLIEESQOQEKNOQELLOLDKWSL 675  
RESULT 11  
VCLJKB  
env polyprotein precursor - human immunodeficiency virus type 1 (strain KB-1-gp41)  
N:Alternate names: coat polyprotein  
N:Contains: coat protein gp120; coat protein gp41  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 01-Mar-1996  
C:Accession: A42995  
R:Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura, T.  
Virology 189, 534-546, 1992  
A:Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated  
A:Reference number: A42995; MUID:92351552; PMID:1322587  
A:Accession: A42995  
A:Molecule type: mRNA  
A:Residues: 1-861 <SHI>  
A:Cross-references: GB:S41266; GB:D01206  
C:Genetics:  
C:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
F:1-689/Domain: extracellular #status predicted <EXT>  
F:1-33/Domain: signal sequence #status predicted <SIG>  
F:17-33/Region: hydrophobic #status predicted  
F:34-517/Product: coat protein gp120 #status predicted <CP1>  
F:514-517/Region: cleavage processing #status predicted  
F:518-861/Product: coat protein gp41 #status predicted <CP2>  
F:518-534/Region: hydrophobic #status predicted  
F:690-711/Domain: transmembrane #status predicted <TM1>  
F:712-861/Domain: intracellular #status predicted <INT>  
F:756-772/Region: hydrophobic #status predicted  
F:793,141,145,146,163,191,192,237,241,248,269,283,296,308,338,345,361,367,397,403,408,414

Query Match 80.8%; Score 2808; DB 1; Length 861;  
Best Local Similarity 81.8%; Pred. No. 2e-198;  
Matches 534; Conservative 49; Mismatches 48; Indels 22; Gaps 9;  
Qy 3 LWTVYGVVPWKATTLFCASDAKAYDTEAHNVWATHACVPDPNPQEVVLNVTFN 62  
Db 39 LWTVYGVVPWKATTLFCASDAKAYDTEAHNVWATHACVPDPNPQEVVLNVTFN 98  
Qy 63 NMWKNVVEQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNTNTNTTSLIIVVWEQ 122  
Db 99 NMWKNVVEQMHENIISLDQSLKPCVKLTPLCVTLNCTDL--RNTTNNS-SI---EE 151  
Qy 123 RGKGMNCSFNITTSIRDKVQREYALFYKLDVPEIDDNKNNTNTKYRLNCSVTIQ 182  
Db 152 KMKGEIKNCSFNVTNIRDKVQREYALFYKLDVPEIDDNKNST-NTCYRLISCDTSVITQ 210  
Qy 183 ACPKVSPEPIPIHYCTPTGTFALLKCNCKKFGTGPCTNVSTVQCTHGRIPVYSTQLLNG 242  
Db 211 ACPKVSPEPIPIHYCTPAGFALLKCNCKNTFNGTGPCKNVSTVQCTHGRIPVYSTQLLNG 270  
Qy 243 SLAEEVVRSENFNTNNAKTIIVQLNVSVETNCPNHNTRKRVTLGPRVWYTTGILG 302  
Db 271 SLAEEVVRSENFNTDNVKTIIVQLNETVKINCIRPNKTKRVTMGPRVYTTGILG 330  
Qy 303 NROAHCHNISRAQWNTLQOIATTLREQFGNKTIAFNQSOGDDEIIVMHSFNCGGEFFYC 362  
Db 331 DIRAHCHNISRAEWNTLEQIANLKRQFENKTIIVFNQSOGDDEIIVMHSFNCGGEFFYC 390  
Qy 363 NSTOLFNSAMNVTNGTWSVTRKQDGT-DIITLPCRIKQIINWQVGVKAMYPALPIKGL 421  
Db 391 DSSQLENSTH--LSNGTW-----WNGTGPENITLPCRIKQIINWQVGVKAMYPALPIKGL 443  
Qy 422 IRCSSNTIGLLTRDGG-----GENOTTEIFRPGGDMRDNWRSELYKYKVKIEPLGVA 476

Db 444 IRCSSNITGLLLTRDGGNTQNNNTSSIEIFRPGGDMRDNWRSELYKYKVKIEPLGVA 503  
Qy 477 PTKAKRRVQREKRAVGLGAMFLGFGAAGSTMGATSMALTIVQARQLLSGIVQQQNLL 536  
Db 504 PTKAKRRVQREKRAVG-IGAVFLGFLGAAGSTWGAATVLTIVQARQLLPQIVQQQNLL 562  
Qy 537 RAIKAQHLQLTVMGIKQARILAVERYLKDOQLGFGWCSGKLICTTAVPNNASHN 596  
Db 563 RAIDAOQHLQLTVMGIKQARILAVERYLKDOQLGFGWCSGKLICTTAVPNNASHN 622  
Qy 597 KTLQIWNMTWMEWDREIDNTHLYIETIESNQKQKNOELLQDLKWSL 649  
Db 623 KSFNEIWNMTWMEWDREINNTYLIETIESNQKQKNOELLQDLKWSL 675  
RESULT 12  
VCLJH3  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)  
N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03973  
R:Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Dora-  
nberger, J.A.; Papas, T.S.; Grzyb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.  
Nature 313, 277-284, 1985  
A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.  
A:Reference number: A93353; MUID:85111123; PMID:2578615  
A:Accession: A03973  
A:Molecule type: DNA  
A:Residues: 1-856 <RAT>  
A:Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:g326383; PIDN:AAA442  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro-  
tein; 1-30/Domain: signal sequence #status predicted <SIG>  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-511/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406  
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predict  
Query Match 80.8%; Score 2806.5; DB 1; Length 856;  
Best Local Similarity 82.8%; Pred. No. 2.5e-198;  
Matches 539; Conservative 35; Mismatches 58; Indels 19; Gaps 9;  
Qy 3 LWVTYYGVVPWKATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 62  
Db 34 LWVTYYGVVPWKATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 93  
Qy 63 NMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVLTNCTDLNNTNTTSLTSIIVVWEQ 122  
Db 94 NMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVLTNCTDLNNTNTSSSGRMIME- 150  
Qy 123 RGKGMRCNSNITTSIRDKVQREYALFYKLDVPEIDNKNNTNTKYLINCMNTSVITQ 182  
Db 151 --KGEIKNCNSFNISTSRGKVKQKEYAFYKLDIIPDN-----DTTSYLTSCNTSVITQ 203  
Qy 183 ACPKVSPEPIPIHYCTPTGFAALLKCNKNGTGPCTNVSTVQCHGIRPVVSTQLLNG 242  
Db 204 ACPKVSPEPIPIHYCAPAGFAILKCNKNTFNGTGPCTNVSTVQCHGIRPVVSTQLLNG 263  
Qy 243 SLAEEVVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRK--RVTLGPGRVMTTGEI 300  
Db 264 SLAEEVVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRK--RVTLGPGRVMTTGEI 323  
Qy 301 LGNTROAHCHNSRAQWNTLQIATTLREQFG-NKTIAFNQSOGGDPPIVHNSFCGGEF 359  
Db 324 -GNMRAHCNISRKWNNTLKQIDSKLREQFGNNKTIIFKQSSGGDPPIVHNSFCGGEF 382  
Qy 360 FYCNSTQLFNSTW---FNSTWSTKGSNNTSGSDTIITLPCRILQIINMMQEVGKAWYALPI 418

Db 383 FYCNSTQLFNSTW---FNSTWSTKGSNNTSGSDTIITLPCRILQIINMMQEVGKAWYALPI 439  
Qy 419 KGLIRCSNITGLLLTRDGGGENTTEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPT 478  
Db 440 SGQIRCSNITGLLLTRDGGGNSNNEIFRPGGDMRDNWRSELYKYKVKIEPLGVAPT 499  
Qy 479 KAKRRVQREKRAVGLGAMFLGFGAAGSTMGATSMALTIVQARQLLSGIVQQQNLLRA 538  
Db 500 KAKRRVQREKRAVG-IGAVFLGFLGAAGSTMGATSMALTIVQARQLLSGIVQQQNLLRA 558  
Qy 539 IKAQHLQLTVMGIKQARILAVERYLKDOQLGFGWCSGKLICTTAVPNNASHNKT 598  
Db 559 IEAQHLQLTVMGIKQARILAVERYLKDOQLGFGWCSGKLICTTAVPNNASHNKS 618  
Qy 599 LQIWNMTWMEWDREIDNTHLYIETIESNQKQKNOELLQDLKWSL 649  
Db 619 LEQIWNMTWMEWDREINNTYLIETIESNQKQKNOELLQDLKWSL 669  
RESULT 13  
VCLJVL  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LV)  
N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03974  
R:Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.  
Nature 313, 450-458, 1985  
A:Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovi-  
A:Reference number: A93355; MUID:85111157; PMID:2982104  
A:Accession: A03974  
A:Molecule type: DNA  
A:Residues: 1-856 <MUE>  
A:Cross-references: GB:K02083; NID:g555008; PIDN:AAAB59873.1; PID:g328559  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro-  
tein; 1-30/Domain: signal sequence #status predicted <SIG>  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-511/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406,  
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predict  
Query Match 80.6%; Score 2799.5; DB 1; Length 856;  
Best Local Similarity 82.5%; Pred. No. 8.2e-198;  
Matches 537; Conservative 36; Mismatches 59; Indels 19; Gaps 9;  
Qy 3 LWVTYYGVVPWKATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 62  
Db 34 LWVTYYGVVPWKATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTENF 93  
Qy 63 NMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVLTNCTDLNNTNTTSLTSIIVVWEQ 122  
Db 94 NMWKNMVEQMHEDIISLWQSLKPCVKLTPLCVLTNCTDLNNTNTSSSGRMIME- 150  
Qy 123 RGKGMRCNSNITTSIRDKVQREYALFYKLDVPEIDNKNNTNTKYLINCMNTSVITQ 182  
Db 151 --KGEIKNCNSFNISTSRGKVKQKEYAFYKLDIIPDN-----DTTSYLTSCNTSVITQ 203  
Qy 183 ACPKVSPEPIPIHYCTPTGFAALLKCNKNGTGPCTNVSTVQCHGIRPVVSTQLLNG 242  
Db 204 ACPKVSPEPIPIHYCAPAGFAILKCNKNTFNGTGPCTNVSTVQCHGIRPVVSTQLLNG 263  
Qy 243 SLAEEVVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRK--RVTLGPGRVMTTGEI 300  
Db 264 SLAEEVVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRK--RVTLGPGRVMTTGEI 323  
Qy 301 LGNTROAHCHNSRAQWNTLQIATTLREQFG-NKTIAFNQSOGGDPPIVHNSFCGGEF 359  
Db 324 -GNMRAHCNISRKWNNTLKQIDSKLREQFGNNKTIIFKQSSGGDPPIVHNSFCGGEF 382



QY 360 FYCNSTOLFNSAMNVTSGTWSVTRKQDGT-DIITLPCRIKQIINRWQVGVKAMVALPI 418  
 Db 383 FYCNSTOLFNSMTW---FNSTWSTEGSNNTGSDTITLPCRIKQIINRWQVGVKAMVAPPI 439  
 QY 419 KGLIRCSSNITGLLLTRDGGGENTTEIFRPGGDMRDNRSELYKYVVKVIEPLGVAPT 478  
 Db 440 SGQIRCSSNITGLLLTRDGGGNNNESEIFRPGGDMRDNRSELYKYVVKVIEPLGVAPT 499  
 QY 479 KAKRRVVQREKRAVGMFLGFLGAAGTMGATSMALTQVQRLLSGIVQOQNLLRA 538  
 Db 500 KAKRRVVQREKRAVG-IGALFLGFLGAAGTMGAASMTLTQVQRLLSGIVQOQNLLRA 558  
 QY 539 IKAQOHLQLTQVWGIKQIARILAVERYLKDQQLGFWGCGSKLICCTTAVPWNASNKT 598  
 Db 559 IEAQOHLQLTQVWGIKQIARILAVERYLKDQQLGFWGCGSKLICCTTAVPWNASNKS 618  
 QY 599 LDGIWNNMTWMDREIDNYTHLIYTLIERSONQOEKNOQELQLDKWASL 649  
 Db 619 LEQIWNNTWMDREIDNYTHLSLIERSONQOEKNEQELLELDKWANL 669  
 RESULT 14  
 VCLJLV  
 env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)  
 N:Alternate names: coat polyprotein  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 A:Note: Host Homo sapiens (man)  
 C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
 C:Accession: A03975  
 R:Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.  
 Cell 40, 9-17, 1985  
 A:Title: Nucleotide sequence of the AIDS virus, LAV.  
 A:Reference number: A90866; MUID:85099333; PMID:2981635  
 A:Accession: A03975  
 A:Molecule type: DNA  
 A:Residues: 1-861 <WAI>  
 A:Cross-references: GB:K02013; NID:g326417; PIDN:AA859751.1; PID:g326424  
 C:Genetics:  
 A:Gene: env  
 C:Superfamily: type E retrovirus env polyprotein  
 C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot  
 F:1-30/Domain: signal sequence #status predicted <SIG>  
 F:31-516/Product: exterior membrane glycoprotein #status predicted <EXT>  
 F:517-861/Product: transmembrane glycoprotein #status predicted <TM>  
 F:88,136,141,146,161,165,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411  
 F:616,621,630,642,679,755,821/Binding site: carboxydrate (Asn) (covalent) #status predic  
 Query Match 80.5%; Score 2795; DB 1; Length 861;  
 Best Local Similarity 82.4%; Pred. No. 1.8e-197;  
 Matches 539; Conservative 37; Mismatches 58; Indels 20; Gaps 10;  
 QY 3 LWTVYGVVWKEATTLFCASDAKAYDEAHNWAHACVPTNPQVWLVNVTENF 62  
 Db 34 LWTVYGVVWKEATTLFCASDAKAYDEVHNWATHACVPTDNPQVWLVNVTENF 93  
 QY 63 NMKNMNVQMHEDIISLWDQSLKPCVKLTPLCVLTNCTDIN--TN--NTNTTSLTIV 119  
 Db 94 NMKNMNVQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLGNATNNTSSNGSEM 153  
 QY 120 WEQKGEMRNCNFTTTSIRDKVQREYALFYKLDVEFDIDNKNNTNNTKYRLINCNTSV 179  
 Db 154 ME---KGEIKNCNFTSIRGVQKEYAFYKLDIIPIDN----DTTSYLTSCNTSV 205  
 QY 180 ITQACPKVSEPIPIHYCTGTGALLKCNKNGTGPCTNVSTVQCTHGIRPWSVQTL 239  
 Db 206 ITQACPKVSEPIPIHYCAPAGEFALLKCNKNTGTGCTNVSTVQCTHGIRPWSVQTL 265  
 QY 240 LNSLAEEVVRSENFNTNAKTIIVQLNVSVEINCTRPNNHTRK--RVTLGPRVWYTT 297  
 Db 266 LNSLAEEVVRSENFNTNAKTIIVQLNVSVEINCTRPNNHTRKIRIRGPGRAFTI 325  
 QY 298 GEILGNIRQAHNCISRAQWNNTLQIATTLREQFG-NKTIAPNQSOGDPEIVHSHFNCG 356

Db 326 GKI-GNMROAHNCISRAQWNNTLQIATTLREQFGNKNKTIIFKQSSGGDPEIVHSHFNCG 384  
 QY 357 GEFFYCNSTOLFNSAMNVTSGTWSVTRKQDGT-DIITLPCRIKQIINRWQVGVKAMYA 415  
 Db 385 GEFFYCNSTOLFNSMTW---FNSTWSTEGSNNTGSDTITLPCRIKQIINRWQVGVKAMYA 441  
 QY 416 LPTKGLIRCSSNITGLLLTRDGGGENTTEIFRPGGDMRDNRSELYKYVVKVIEPLGV 475  
 Db 442 PPSQIRCSSNITGLLLTRDGGGNNNGSEIFRPGGDMRDNRSELYKYVVKVIEPLGV 501  
 QY 476 APTKAKRRVVQREKRAVGMFLGFLGAAGTMGATSMALTQVQRLLSGIVQOQNLL 535  
 Db 502 APTKAKRRVVQREKRAVG-IGALFLGFLGAAGTMGAASMTLTQVQRLLSGIVQOQNLL 560  
 QY 536 LRAIKAQOHLQLTQVWGIKQIARILAVERYLKDQQLGFWGCGSKLICCTTAVPWNASWS 595  
 Db 561 LRAIEAQOHLQLTQVWGIKQIARILAVERYLKDQQLGFWGCGSKLICCTTAVPWNASWS 620  
 QY 596 NKTLDQIWNNTWMDREIDNYTHLIYTLIERSONQOEKNOQELQLDKWASL 649  
 Db 621 NKSLEQIWNNTWMDREIDNYTHLSLIERSONQOEKNEQELLELDKWASL 674  
 RESULT 15  
 S33985  
 env polyprotein - human immunodeficiency virus type 1  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 26-Aug-1999  
 C:Accession: S33985  
 R:Carlini, F.  
 submitted to the EMBL Data Library, November 1991  
 A:Reference number: S33979  
 A:Accession: S33985  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-851 <CAR>  
 A:Cross-references: EMBL:Z11530; NID:g60192; PIDN:CAA77628.1; PID:g60199  
 C:Superfamily: type E retrovirus env polyprotein  
 Query Match 80.1%; Score 2782; DB 2; Length 851;  
 Best Local Similarity 82.0%; Pred. No. 1.6e-196;  
 Matches 534; Conservative 39; Mismatches 54; Indels 24; Gaps 9;  
 QY 3 LWTVYGVVWKEATTLFCASDAKAYDEAHNWAHACVPTNPQVWLVNVTENF 62  
 Db 34 LWTVYGVVWKEATTLFCASDAKAYDEVHNWATHACVPTDNPQVWLVNVTENF 93  
 QY 63 NMKNMNVQMHEDIISLWDQSLKPCVKLTPLCVLTNCTDINNTNTTNTTSLTIVVWEQ 122  
 Db 94 NMKNMNVQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDL--KNDINTNSSSGRIME- 150  
 QY 123 RGKGMWRNCNFTTTSIRDKVQREYALFYKLDVEFDIDNKNNTNNTKYRLINCNTSVITQ 182  
 Db 151 --KGEIKNCNFTSIRGVQKEYAFYKLDIIPIDN----DTTSYLTSCNTSVITQ 203  
 QY 183 ACPKVSFEPIPIHYCTGTGALLKCNKNGTGPCTNVSTVQCTHGIRPWSVQTL 242  
 Db 204 ACPKVSFEPIPIHYCAPAGEFALLKCNKNTGTGCTNVSTVQCTHGIRPWSVQTL 263  
 QY 243 SLAEEVVRSENFNTNAKTIIVQLNVSVEINCTRPNNHTRKVTL--GPRVWYTTGEI 300  
 Db 264 SLAEEVVRSENFNTNAKTIIVQLNVSVEINCTRPNNHTRKIRIRGPGRAFTIGKI 323  
 QY 301 LGNIRQAHNCISRAQWNNTLQIATTLREQFG-NKTIAPNQSOGDPEIVHSHFNCGGEF 359  
 Db 324 -GNMROAHNCISRAQWNNTLQIATTLREQFGNKNKTIIFKQSSGGDPEIVHSHFNCGGEF 382  
 QY 360 FYCNSTOLFNSAMNVTSGTWSVTRKQDGT-DIITLPCRIKQIINRWQVGVKAMVALPI 418  
 Db 383 FYCNSTOLFNSMTWSTKGSNT-----EGSTITLPCRIKQIINRWQVGVKAMVAPPI 434  
 QY 419 KGLIRCSSNITGLLLTRDGGGENTTEIFRPGGDMRDNRSELYKYVVKVIEPLGVAPT 478



```

Db 435 SGOIRGSSNITGLLTRDGGNSNNESEIFRPGGDMRDNRSELYKYXVVKIEPLGVAPT 494
QY 479 KAKRRVVQREKRAVGM/GAMFLGFLGAAGSTMGATSMALTVOAROLLSGIVQOQNNLLRA 538
Db 495 KAKRRVVQREKRAVG--IGALFLGFLGAAGSTMGAASMTLTVOAROLLSGIVQOQNNLLRA 553
QY 539 IKAQQHLLQLTVWGIKQLQARILAVERYLKDQQLLGFWCSGKLICTTAVPWNASWSNKT 598
Db 554 IEAQOHLQLTVWGIKQLQARILAVERYLKDQQLLGFWCSGKPICTTAVPWNASWSNKS 613
QY 599 LDCIWNNTWMEWDREIDNYTHIYTLIEESQOQEKNOQELLQLDKWASL 649
Db 614 LEQIWNNTWMEWDREINNYTSLIHSLEESQOQEKNEQELLELDKWASL 664

```

Search completed: January 13, 2004, 09:11:58  
Job time : 23 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2004, 08:59:06 ; Search time 48 Seconds

(without alignments)  
2146.116 Million cell updates/sec

Title: US-09-938-406-1\_COPY\_33\_681

Perfect score: 3474

Sequence: 1 ANLWTVYVGVVWKEATT.....QNQKEKNQELLQDKWASL 649

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3457	99.5	649	19 AAW44250	HIV-1 gp160 residu
2	3388	97.5	865	19 AAW43072	HIV-1 gp120 protei
3	2977.5	85.7	883	22 AAB82761	Ancestral HIV-1 gr
4	2883.5	83.0	847	21 AAY97073	Variant HIV-1 SF16
5	2878.5	82.9	842	24 ABU66565	Human immunodefici
6	2870.5	82.6	842	23 ABB06211	HIV Env isolate SF
7	2852	82.1	850	16 AAR67724	gp120 from the HIV
8	2834	81.6	855	18 AAW11581	Human Immunodefici
9	2834	81.6	855	20 AAW88113	Env protein of the

HIV-1 SOS gp140 gl  
HIV protein Hf6.  
ENV protein contai  
HIV-1 (ATCC CRL 85  
HIV-1 (ATCC CRL 85  
AIDS associated re  
AIDS associated re  
Sequence of ARV-2  
env gene decoded f  
HIV-1 env protein.  
HIV-1 coat protei  
HIV gp41 DNA SEQ I  
HIV-1 NL4-3 gp120  
Human immunodefici  
HIV envelope prote  
HIV envelope prote  
HIV-1 strain OVI E  
Protein encoded by  
Wild type HIV-1 HX  
gp120 from the HIV  
Amino acid sequenc  
HIV-III ENV-LOR g  
HIV env protein fr  
HIV-SF2 virus gp12  
Sequence of envelo  
HIV-1 env protein  
Modified HIV env g  
HIV-III virus (HI  
HIV-1/IIIB env clo  
Lymphadenopathy-as  
Sequence of LAV vi  
HIV-1 env segment  
gp120 from the HIV  
Sequence of the AI  
HIV protein HT7.  
HIV virus env gene

#### ALIGNMENTS

##### RESULT 1

AAW44250  
ID AAW44250 standard; protein; 649 AA.

XX AAW44250;

XX 26-JUN-1998 (first entry)

XX HIV-1 gp160 residues 33-681.

DE Vaccine; antibody; antigen; hydrophobic; proteosome; pathogen;  
KW immune response; sexually transmitted disease; HIV; infection.

OS Human immunodeficiency virus type 1.

XX Key Location/Qualifiers

FT Region 491..519  
FT /note= "hydrophobic region"

FN WO9801558-A2.

XX 15-JAN-1998.

XX 10-JUL-1997; 97WO-US12253.

XX 10-JUL-1996; 96US-0021687.

XX (INTE-) INTELLIVAX INC.

PA (JACK-) JACKSON FOUND HENRY M.

XX (USSA) US SEC OF ARMY.

PI Birx DL, Lowell GH, Vancott TC;

XX WPI; 1998-110231/10.

XX Vaccine compositions for eliciting neutralising antibodies -

XX comprising antigen containing hydrophobic sequence or having added

XX hydrophobic material, complexed to proteosomes or bio-adhesive

XX nano-emulsions

XX Claim 9; Page 22; 62pp; English.

XX The present sequence represents HIV-1 gp160 residues 33-681 (the full

XX protein is on the SWISS-PROT database Seq ID: 119434) used in a

XX vaccine of the present invention. The vaccine composition is capable of

XX eliciting neutralising antibodies in a subject to a pathogenic organism

XX which antibodies are present in vaginal secretions, intestinal

XX secretions, lung secretions or faeces. The vaccine comprises: (a) an

XX antigen comprising a protein or peptide having: (i) an endogenous

XX hydrophobic sequence of 3-50 non-polar or uncharged amino acids; (ii)

XX added to the protein or peptide, an exogenous hydrophobic material

XX comprising a sequence of 3 to 50 non-polar or uncharged amino acids or a

XX 8-18C fatty acyl group, or (iii) both (i) and (ii), and (b) complexed

XX with the antigen, a composition comprising proteosomes, bioadhesive

XX nano-emulsions or both, where the complexed or coupled protein or

XX peptide maintains a native structure of antigenic epitopes such that,

XX upon administration to the subject, the antigen induces neutralising

XX antibodies in one or more of vaginal secretions, intestinal secretions,

XX lung secretions and faeces, capable of neutralising the pathogenic

XX organism. The compositions can be used for inducing an immune response

XX against a pathogenic organism such as a causative agent of a sexually-

XX transmitted or mucosally-transmitted disease, e.g. HIV infection. The

XX compositions preserve the antigenic integrity of the protein or peptide

XX epitopes while at the same time enhancing their immunogenicity.

XX Sequence. 649 AA;

Query Match 99.5%; Score 3457; DB 19; Length 649;

Best Local Similarity 99.5%; Pred. No. 4.3e-187;

Matches 646; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ANLWTVVYGVVPMKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTE 60

DB 1 ANLWTVVYGVVPMKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTE 60

QY 61 NFNWKNVNMVQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTNTTSLIIVW 120

DB 61 NFNWKNVNMVQMHEDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTNTTSLIIVW 120

QY 121 EORKGEMRNCSENIITTSIRDKVQREYALFYKLDVEPTDDNKNTNTNTKYRLINCNTSVI 180

DB 121 EORKGEMRNCSENIITTSIRDKVQREYALFYKLDVEPTDDNKNTNTNTKYRLINCNTSVI 180

QY 181 TQACPKVSFEPIPIHYCTPTGFALLKCNCKXFNCTVSTVQCTHGRPVVSTQLLL 240

DB 181 TQACPKVSFEPIPIHYCTPTGFALLKCNCKXFNCTVSTVQCTHGRPVVSTQLLL 240

QY 241 NGSIAEEVWIRSENFNTNAKTIIVQLNVSEINCTRNNTNRKRVTLGPGRVWTTGFI 300

DB 241 NGSIAEEVWIRSENFNTNAKTIIVQLNVSEINCTRNNTNRKRVTLGPGRVWTTGFI 300

QY 301 LGNIRQAHNCISRAQWNNTLQOIAATLRFQGNKTIATNQSNGGDPETVWHSFNGCGEFFF 360

DB 301 LGNIRQAHNCISRAQWNNTLQOIAATLRFQGNKTIATNQSNGGDPETVWHSFNGCGEFFF 360

QY 361 YCNSTQLENSAWNVTSNGTWSVTRKQDGTGDIITLPCRIKQIINRWQVVGKAMYPALPIK 420

DB 361 YCNSTQLENSAWNVTSNGTWSVTRKQDGTGDIITLPCRIKQIINRWQVVGKAMYPALPIK 420

QY 421 LIRCSSNTTGLLLTRDGGENOTTBIFRPGGDMRDNWRSLEYKYVKVYKIBPLGVAPTKA 480

DB 421 LIRCSSNTTGLLLTRDGGENOTTBIFRPGGDMRDNWRSLEYKYVKVYKIBPLGVAPTKA 480

QY 481 KRRVVOREKRAVGMGLGAMFLGFLGAGSTMGATSMALTVOARQLLSGIVQQNNLLRAIK 540

DB 481 KRRVVOREKRAVGMGLGAMFLGFLGAGSTMGATSMALTVOARQLLSGIVQQNNLLRAIK 540

QY 541 AQOHLIQLTVMGIKOLQARILAVERYLKQOOLGFGWCSGKLICTTAVPWNASWNTLTD 600

DB 541 AQOHLIQLTVMGIKOLQARILAVERYLKQOOLGFGWCSGKLICTTAVPWNASWNTLTD 600

QY 601 QIWNNTWMEWDREIDNTHYLIYTLIEESQNOQEKNOQELLQLDKWSL 649

DB 601 QIWNNTWMEWDREIDNTHYLIYTLIEESQNOQEKNOQELLQLDKWSL 649

## RESULT 2

AAW43072

ID AAW43072 standard; peptide; 865 AA.

XX AAW43072;

AC AAW43072;

XX 11-SEP-1998 (first entry)

XX HIV-1 gp120 protein fragment from isolate CDC4.

XX gp120 protein; purification; fractionation; ion exchange; chromatography;

XX binding affinity; CD4; hydrophobic interaction; size exclusion; vaccine.

XX Human immunodeficiency virus type 1.

XX US5696238-A.

XX 09-DEC-1997.

XX 11-MAY-1995; 95US-0439286.

XX 20-AUG-1991; 91US-0684963.

XX 16-AUG-1993; 93US-0109002.

XX 09-MAY-1994; 94US-0240073.

XX 11-MAY-1995; 95US-0439286.

XX (CHIR ) CHIRON CORP.

XX Haigwood NL, Scandella C;

XX WPI; 1998-041353/04.

XX Purification of HIV gp120 - using chromatographic methods

XX Disclosure; Fig 2A-W; 53pp; English.

XX AAW43066-W43080 are fragments of the gp120 protein from different human

XX immunodeficiency virus type I (HIV-1) isolates. These proteins are used

XX in a novel method for purifying HIV gp120 so as to provide a purified

XX gp120 glycoprotein having protein/protein binding properties

XX substantially identical to natural viral HIV gp120. The method involves

XX fractionating a crude gp120 preparation containing full-length,

XX glycosylated gp120 using ion exchange chromatography so as to provide a

XX first collection of fractions. A fraction from the first collection is

XX selected that exhibits specific binding affinity for CD4 peptide,

XX thereby producing a first fractionated material. The first fractionated

XX material is fractionated by hydrophobic interaction chromatography so as

XX to provide a second collection of fractions from which a second

XX collection is selected that exhibits specific binding affinity for CD4

XX peptide. This second fraction is fractionated by size exclusion

XX chromatography so as to provide a third collection of fractions

XX exhibiting specific binding affinity for CD4 peptide, thereby providing

XX the purified gp120. The purified gp120 can be used for antibody

XX production and in vaccines.

XX Sequence 865 AA;

Query Match 97.5%; Score 3388; DB 19; Length 865;

Best Local Similarity 98.9%; Pred. No. 4.2e-183;

Matches 644; Conservative 1; Mismatches 0; Indels 6; Gaps 6;

QY 1 ANLWTVVYGVVPMKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTE 60

Db 32 ANLWTVYGVPKKEA-TTLFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTE 90  
 QY 61 NFNMMKNMVEQMHEDIISLWDQSLKPCVKLPCLVTLNCTDLNNTTN-TTSLIIV 119  
 Db 91 NFNMMKNMVEQMHEDIISLWDQSLKPCVKLPCLVTLNCTDLNNTTN-TTSLIIV 150  
 QY 120 WEORGKGMRCSENIITTSIRDKVQREYALFYKLDVEPIDDKNK-TTNTKYLINCN 178  
 Db 151 WEORGKGMRCSENIITTSIRDKVQREYALFYKLDVEPIDDKNK-TTNTKYLINCN 210  
 QY 179 VITQACPKVSFEPIPIHYCTPTGFALLKCNKKFNGTGTCTNVSTVQCTHGRPVVSTOL 238  
 Db 211 VITQACPKVSFEPIPIHYCTPTGFALLKCNKKFNGTGTCTNVSTVQCTHGRPVVSTOL 270  
 QY 239 LLNGSLAEVEVIRSENFNNAKTIIVQLNVSVNEINCTPNNHTRKRVTLGPRVWYTG 298  
 Db 271 LLNGSLAEVEVIRSENFNNAKTIIVQLNVSVNEINCTPNNHTRKRVTLGPRVWYTG 330  
 QY 299 EILGNIRQAHONISRAQWNTLQQIATTLREQFGNKTIAFNOSGGDPEIVMHSFNC 358  
 Db 331 QILGNIRQAHONISRAQWNTLQQIATTLREQFGNKTIAFNOSGGDPEIVMHSFNC 390  
 QY 359 FFCNSTOLFNSAMNVTSGTWSVRKQKDTGDIITLPCRIKQIINRWGVGKAMVALPI 418  
 Db 391 -FCNSTOLFNSAMNVTSGTWSVRKQKDTGDIITLPCRIKQ-INRWGVGKAMVALPI 448  
 QY 419 KGLIRCSSNITGLLTRDGGNGQTTEIPRPGGDMRDNRSELYKYKVKIEPIGVAPT 478  
 Db 449 KGLIRCSSNITGLLTRDGGNGQTTEIPRPGGDMRDNRSELYKYKVKIEPIGVAPT 508  
 QY 479 KAKRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQQNNLLRA 538  
 Db 509 KAKRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQQNNLLRA 567  
 QY 539 IKAQQLHLQLTWGKIKQQLARILAVERYLKQQLLGFWGCSEKLICTTAVPNWASWSNKT 598  
 Db 568 IKAQQLHLQLTWGKIKQQLARILAVERYLKQQLLGFWGCSEKLICTTAVPNWASWSNKT 627  
 QY 599 LQIWNNTWMEWDREIDNYTHLYTLIEESQKQKQQLLQDKWASL 649  
 Db 628 LQIWNNTWMEWDREIDNYTHLYTLIEESQKQKQQLLQDKWASL 678

## RESULT 3

AAB82761  
 ID AAB82761 standard; Protein; 883 AA.

AC AAB82761;

DT 29-OCT-2001 (first entry)

XX Ancestral HIV-1 group M, subtype B gp160 protein.

DE HIV-1; env gene; gp160 protein; ancestral gene; phylogeny; vaccine.

KW Human immunodeficiency virus type 1.

OS WO200160838-A2.

FN 23-AUG-2001.

PD 16-FEB-2001; 2001WO-US05288.

PF 18-FEB-2000; 2000US-0183659.

XX (UNIW ) UNIV WASHINGTON.

PA Mullins JI, Rodrigo AG, Learn GH, Li F;

PI WPI; 2001-536565/59.

DR N-PSDB; AAB26468.

XX

PT Preparing an ancestral viral amino acid sequence useful as a vaccine  
 PT comprises determining a recent ancestor of a circulating virus by  
 maximum likelihood phylogeny analysis -

Claim 8; Page 54; 89pp; English.

CC The present sequence is that of an ancestral HIV-1 group M, subtype  
 CC B gp160 (env gene product) sequence. The invention provides  
 CC compositions and methods for determining ancestral viral gene  
 CC sequences and ancestral viral protein sequences for highly diverse  
 CC viruses, such as HIV-1. The methods use samples of circulating  
 CC viruses to determine an ancestral viral sequence by maximum  
 CC likelihood phylogeny analysis. In the present case, the ancestral  
 CC HIV-1 subtype B env sequence (see AAB26468) was determined using 38  
 CC subtype B sequences (obtained from 9 different countries) and 3  
 CC subtype D (outgroup) sequences. The distances between this ancestral  
 CC viral sequence and circulating strains used to determine it were on  
 CC average 12.3% (range: 8.0-21.0%) while the available specimens were  
 CC 17.3% different from each other (range: 13.3-23.2%). Thus, the  
 CC ancestor sequence was, on average, more closely related to any given  
 CC circulating virus than to any other variant. The ancestral gp160  
 CC sequence included a wide variety of immunogenically active peptides  
 CC when processed for antigen presentation; nearly all known subtype B  
 CC CTL epitope consensus amino acids were represented. Thus, an  
 CC immunogenic composition to this subtype B ancestor protein will  
 CC elicit broad neutralising antibody against HIV-1 isolates of the  
 CC same subtype, and will also elicit a broad cellular response  
 CC mediated by antigen-specific T-cells. A claimed vaccine composition  
 CC comprises a viral ancestor protein or its immunogenic fragment,  
 CC especially one derived from the HIV-1 group M subtype B gp160  
 CC ancestral protein.

SQ Sequence 883 AA;

Query Match 85.7%; Score 2977.5; DB 22; Length 883;

Best Local Similarity 84.7%; Pred. No. 5.7e-160;

Matches 564; Conservative 27; Mismatches 54; Indels 21; Gaps 6;

QY 3 LWTVYGVVPMKEATTLFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTE 62

Db 33 LWTVYGVVPMKEATTLFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTE 92

QY 63 NFNMMKNMVEQMHEDIISLWDQSLKPCVKLPCLVTLNCTDLNNTTN-TTSLIIV 122

Db 93 NFNMMKNMVEQMHEDIISLWDQSLKPCVKLPCLVTLNCTDLNNTTN-TTSLIIV 152

QY 123 RG-----KGEMNCSENIITTSIRDKVQREYALFYKLDVEPID-DNKNTNTNTKYRLIN 174

Db 153 SGGGTMEGEKGEIKNCSFNVTTSIRDKVQREYALFYKLDVEPID-DNKNTNTNTKYRLIN 212

QY 175 QNTSVITQACPKVSFEPIPIHYCTPTGFALLKCNKKFNGTGTCTNVSTVQCTHGRPVV 234

Db 213 QNTSVITQACPKVSFEPIPIHYCTPTGFALLKCNKKFNGTGTCTNVSTVQCTHGRPVV 272

QY 235 STOLLNGSLAEVEVIRSENFNNAKTIIVQLNVSVNEINCTPNNHTRKRVTLGPRVW 294

Db 273 STOLLNGSLAEVEVIRSENFNNAKTIIVQLNVSVNEINCTPNNHTRKRVTLGPRVW 332

QY 295 YTTGEILGNIRQAHONISRAQWNTLQQIATTLREQFGNK--TIAFNOSGGDPEIVMHS 352

Db 333 YATGKIIGDIRQAHONISRAQWNTLQQIATTLREQFGNKTTIVFNQSSGGDPEIVMHS 392

QY 353 FNCGGEFFYCNSTOLFNSAMNVTSGTWSVRKQKDTGDIITLPCRIKQIINRW 407

Db 393 FNCGGEFFYCNSTOLFNSAMNVTSGTWSVRKQKDTGDIITLPCRIKQIINRW 450

QY 408 VVGKAMVALPIKGLIRCSSNITGLLTRDGGNGQT-----TEIFRPGGDMRDNRSELY 463

Db 451 EVGKAMVALPIKGLIRCSSNITGLLTRDGGNGQT-----TEIFRPGGDMRDNRSELY 510

QY 464 KYKVVKIEPIGVAPTAKRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQ 523

Db 511 KYKVVKIEPIGVAPTAKRRVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQ 570

QY 524 LLSGIVQQNNLLRAIKAAQHLQLTWGKIQARILAVERYIKDQQLLGFWGCSSKLI 583  
 DB 571 LLSGIVQQNNLLRAIEAQHLQLTWGKIQARILAVERYIKDQQLLGFWGCSSKLI 630  
 QY 584 CTTAVPNAWSNKTLDQIWNMTWMEWDEIDNYTHLYTLIEESQOQKNEQELQL 643  
 DB 631 CTTAVPNAWSNKSLLDKIWNMTWMEWDEIDNYTHLYTLIEESQOQKNEQELQL 690  
 QY 644 DKWASL 649  
 DB 691 DKWASL 696

## RESULT 4

AAAY97073  
 ID AAY97073 standard; protein; 847 AA.

XX AC AAY97073;

XX DT 31-OCT-2000 (first entry)

XX DE Variant HIV-1 SF162 Env gp160.

XX KW HIV-1; vaccine; anti-HIV; immunogenic; modified; envelope; Env;  
 KW CD4 binding region; VI/V2 loop; bridging sheet.

XX OS Human immunodeficiency virus type 1 isolate SF162.

XX PN WO200039303-A2.

XX PD 06-JUL-2000.

XX PF 30-DEC-1999; 99WO-US31272.

XX PR 31-DEC-1998; 98US-0114495.

XX PR 29-SEP-1999; 99US-0156670.

XX PA (CHIR) CHIRON CORP.

XX PI Barnett S, Hartog K, Martin E;

XX DR WPI; 2000-465745/40.

XX PT Novel modified HIV Env polypeptides useful as immunizing agents and for  
 PT preparing a vaccine to elicit an immune response against a broad range  
 PT of HIV subtypes

XX PS Claim 5; Page 115-117; 139pp; English.

XX CC Novel immunogenic modified human immunodeficiency virus (HIV) envelope  
 CC (Env) polypeptides having an amino acid deleted or replaced in the region  
 CC corresponding to residues 420-436 or 119-123 and 199-210 relative to  
 CC isolate HXB-2 are disclosed. The modified Env polypeptide is based on HIV  
 CC strain SF162, with numbering relative to isolate HXB-2. The Env  
 CC polypeptides are modified so as to expose at least part of the CD4  
 CC binding region. The modified HIV Env polypeptides, coding polynucleotides  
 CC and constructs, further comprising an adjuvant, are used for inducing an  
 CC immune response in an individual. The method involves administering a  
 CC first composition comprising a polynucleotide encoding the Env  
 CC polypeptide in a priming step and administering a second composition  
 CC comprising a modified Env polypeptide as a booster in an amount  
 CC sufficient to induce an immune response in the individual. The first  
 CC and/or second composition further comprises an adjuvant (claimed). The  
 CC intracellularly produced Env polypeptides can be used for a number of  
 CC diagnostic and therapeutic purposes to determine the presence of reactive  
 CC antibodies/and or Env proteins in a biological sample to aid in the  
 CC diagnosis of HIV infection or disease status or as measure of response to  
 CC immunization.

XX SQ Sequence 847 AA;

Query Match 83.0%; Score 2883.5; DB 21; Length 847;

Best Local Similarity 83.3%; Pred. No. 1.le-154;  
 Matches 544; Conservative 40; Mismatches 38; Indels 31; Gaps 7;

QY 3 LWTVYGVVWKEATTTLCASDAKAYDEAHNVWATHACVPTNPDPQEVVLENTENF 62  
 DB 33 LWTVYGVVWKEATTTLCASDAKAYDEAHNVWATHACVPTNPDPQEVVLENTENF 92  
 QY 63 NMWKNVVEQMHEDIISLWDSLKPCVKLTPLCVLNCNTDNTNTNTTNTTSLIIVWEQ 122  
 DB 93 NMWKNVVEQMHEDIISLWDSLKPCVKLTPLCVLNCNTDNTNTNTTNTTSLIIVWEQ 146  
 QY 123 RGKGEVNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTNTNTNTNTNTNTNTNT 182  
 DB 147 MDRGEIKNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTNTNTNTNTNTNTNTNT 201  
 QY 183 ACPKVSPEPIPIHYCTPTGFALLKCNCKKPNFGPCPTNVSTVQCTHGIRPVVSTQILLNG 242  
 DB 202 ACPKVSPEPIPIHYCAPAGFAILLKCNCKKPNFGPCPTNVSTVQCTHGIRPVVSTQILLNG 261  
 QY 243 SLAEVEVIRSENFNTNAKTIIVQLNVSVSEINCTRPNNHTRKRVTLGPGRVWYTTGELG 302  
 DB 262 SLAEVEVIRSENFNTNAKTIIVQLNVSVSEINCTRPNNHTRKRVTLGPGRVWYTTGELG 321  
 QY 303 NIROAHNCISRAOWNTLQIATTLREQFGNKTIAFNQSSGGDPEIYVHSEFNCGEFFYC 362  
 DB 322 DIROAHNCISGEKWNNTLKQIVTKLOAQFGNKTIIVFKQSSGGDPEIYVHSEFNCGEFFYC 381  
 QY 363 NSTQLFNSANVT-----SNGTSVTRKQDGTITLPCRIKQIINRWQVVGKAMALP 417  
 DB 382 NSTQLFNSANVT-----SNGTSVTRKQDGTITLPCRIKQIINRWQVVGKAMALP 428  
 QY 418 IKGLIRCSSNITGLLITRDGGE-NOTTEIFRPGGDMRDNWRSELYKYVVKTEPLGVA 476  
 DB 429 IRGQIRCSSNITGLLITRDGGE-NOTTEIFRPGGDMRDNWRSELYKYVVKTEPLGVA 488  
 QY 477 PTKAKRRVQREKRAVGMGLAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQQNNLL 536  
 DB 489 PTKAKRRVQREKRAV-TLGNMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQQNNLL 547  
 QY 537 RAIKAAQHLQLTWGKIQARILAVERYIKDQQLLGFWGCSSKLICTTAVPNAWSN 596  
 DB 548 RAIKAAQHLQLTWGKIQARILAVERYIKDQQLLGFWGCSSKLICTTAVPNAWSN 607  
 QY 597 KTDQIWNMTWMEWDEIDNYTHLYTLIEESQOQKNEQELQLDKWASL 649  
 DB 608 KSLDQIWNMTWMEWDEIDNYTHLYTLIEESQOQKNEQELQLDKWASL 660

## RESULT 5

ABU66565  
 ID ABU66565 standard; Protein; 842 AA.

XX AC ABU66565;

XX DT 22-MAY-2003 (first entry)

XX DE Human immunodeficiency virus (HIV) envelope (env) protein #1.

XX KW Immunogenic HIV polypeptide; human immunodeficiency virus; HIV;  
 KW vaccine; gene therapy; packaging cell line; humoral immune response;  
 KW cellular immune response; gene delivery vector; DNA immunisation;  
 KW envelope protein; env.

XX OS Human immunodeficiency virus.

XX PN WO2003004657-A1.

XX PD 16-JAN-2003.

XX PF 05-JUL-2002; 2002WO-US21421.

XX PR 05-JUL-2001; 2001US-303192P.

PR 31-AUG-2001; 2001US-316860P.









CC donated blood, and as an immunogen to raise specific antibodies, for  
 XX HIV-1 diagnosis.

SQ Sequence 855 AA;

Query Match 81.6%; Score 2834; DB 20; Length 855;  
 Best Local Similarity 82.2%; Pred. No. 6.7e-152;  
 Matches 534; Conservative 45; Mismatches 53; Indels 18; Gaps 6;

QY 3 LWTVTYGVVWKKEATTTLCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62  
 DB 33 LWTVTYGVVWKKEATTTLCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 92  
 QY 63 NMKNMVEQHEHDIISLWDQSLKPCVKLTPLCVTLNCTDLN--DN--NTNTTSLIIV 119  
 DB 93 NMKNMVEQHEHDIISLWDQSLKPCVKLTPLCVTLNCTDLNATNGDNTTSSRGKV 152  
 QY 120 WEQRKGEMRNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNNTKYRLINCVTSV 179  
 DB 153 ---GGGEMKNCSEFNITTSIRDKVQREYALFYKLDIAPIDNN---SNNRYELISCVTSV 204  
 QY 180 ITQACPKVSEPIPIHYCTPTGFPALLKNDKFNCTGCTNVSTVQCTHGIRPVVSTOLL 239  
 DB 205 ITQACPKVSEPIPIHYCAPAGFAILCKDKKFGKGPCTNVSTVQCTHGIRPVVSTOLL 264  
 QY 240 LNSGLAEEVIRSENFNTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPGRVWYTTGE 299  
 DB 265 LNSGLAEEVIRSENFNTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPGRVWYTTGE 324  
 QY 300 ILGNTRQAHNCISRAQWNTLQOIAATTLREQFGNKTIAFNOSGGDPEIVMHSFNCGEF 359  
 DB 325 IIGDIRQAHNCISRAQWNTLQOIAATTLREQFGNKTIAFNOSGGDPEIVMHSFNCGEF 384  
 QY 360 FYCNSTQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRILKQIINRWQVGVKAMVALPK 419  
 DB 385 FYCNSTQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRILKQIINRWQVGVKAMVALPK 438  
 QY 420 GLIRCSSNITGLLTRDGGGENTTEIFRPGGDMRDNRSLYKVKVVKIEPLGVAPT 479  
 DB 439 GQIRCSSNITGLLTRDGGGENTTEIFRPGGDMRDNRSLYKVKVVKIEPLGVAPT 498  
 QY 480 AKRRVQREKRAVGMGLGAMFLGAGSTMGATSMALTIVQARQLLSGIVQOQNLLRAI 539  
 DB 499 AKRRVQREKRAVGMGLGAMFLGAGSTMGATSMALTIVQARQLLSGIVQOQNLLRAI 557  
 QY 540 KAOHLLQITVWGIKQOARILAVERYLKDQQLLGFWGCSGKLICTTAVPWNASNKTL 599  
 DB 558 EAOHLLQITVWGIKQOARILAVERYLKDQQLLGFWGCSGKLICTTAVPWNASNKTL 617  
 QY 600 DQIWNNTWMEWDREIDNTYLIYTLIBESQOQKQELQLDQKWSL 649  
 DB 618 NKIWNNTWMEWDREIDNTYLIYTLIBESQOQKQELQLDQKWSL 667

RESULT 10

AAB61505  
 ID AAB61505 standard; Protein; 643 AA.

AC AAB61505;

XX 05-APR-2001 (first entry)

DE HIV-1 SOS gp140 glycoprotein.

KW gp120; anti-HIV; gene therapy; vaccine; viral envelope protein; gp41;  
 KW gp140.

OS Human immunodeficiency virus type 1.

XX WO200100648-A1.

PD 04-JAN-2001.

XX

PF 23-JUN-2000; 2000WO-US17267.

PR 25-JUN-1999; 99US-0340992.

XX (PROG-) PROGENICS PHARM INC.

PA (AARO-) AARON DIAMOND AIDS RES CENT.

XX Binley JM, Schuelke N, Olson WC, Maddon PJ, Moore JP;

XX WPI; 2001-122993/13.

DR N-PSDB; AAF28581.

XX New viral envelope proteins, useful for producing vaccines to treat  
 PT human immunodeficiency virus-1 infections, comprises amino acid  
 PT sequence mutations such that viral transmembrane-surface protein  
 PT complex is more stable

XX Disclosure; Fig 13; 109pp; English.

XX The present invention relates to a viral envelope protein. The viral  
 CC envelope protein comprises a viral surface protein (e.g. glycoprotein  
 CC gp120) and a corresponding viral transmembrane protein (e.g. gp41), in  
 CC which the viral envelope protein contains one or more amino acid sequence  
 CC mutations that enhance the stability of the complex formed between the  
 CC viral surface and transmembrane proteins. The viral envelope protein can  
 CC be used in the treatment of viral infection e.g. HIV-1 infection. The  
 CC present sequence is HIV-1 SOS gp140, which was used in the present  
 CC invention.

SQ Sequence 643 AA;

Query Match 81.4%; Score 2827.5; DB 22; Length 643;  
 Best Local Similarity 82.2%; Pred. No. 1.2e-151;  
 Matches 532; Conservative 40; Mismatches 56; Indels 19; Gaps 4;

QY 3 LWTVTYGVVWKKEATTTLCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62  
 DB 4 LWTVTYGVVWKKEATTTLCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 63  
 QY 63 NMKNMVEQHEHDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTSLIIVWEQ 122  
 DB 64 NMKNMVEQHEHDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTSLIIVWEQ 117  
 QY 123 RGKEMRNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNNTKYRLINCVTSV 182  
 DB 118 -ERGEIKNCSEFNITTSIRDKVQREYALFYKLDVEPIDN----NNTSRLISCVTSV 171  
 QY 183 ACPKVSFPIPIHYCTPTGFPALLKNDKFNCTGCTNVSTVQCTHGIRPVVSTOLLNG 242  
 DB 172 ACPKVSFPIPIHYCAPAGFAILCKDKKFGKGPCKNVSTVQCTHGIRPVVSTOLLNG 231  
 QY 243 SLAEEVIRSENFNTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPGRVWYTTGEILG 302  
 DB 232 SLAEEVIRSENFNTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPGRVWYTTGEILG 291  
 QY 303 NIROAHNCISRAQWNTLQOIAATTLREQFGNKTIAFNOSGGDPEIVMHSFNCGEFFYC 362  
 DB 292 DIRQAHNCISRAQWNTLQOIAATTLREQFGNKTIAFNOSGGDPEIVMHSFNCGEFFYC 351  
 QY 363 NSTQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRILKQIINRWQVGVKAMVALPKGI 422  
 DB 352 NSTQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRILKQIINRWQVGVKAMVALPKGI 405  
 QY 423 RGSNITGLLTRDGGGENTTEIFRPGGDMRDNRSLYKVKVVKIEPLGVAPTAKR 482  
 DB 406 RGSNITGLLTRDGGGENTTEIFRPGGDMRDNRSLYKVKVVKIEPLGVAPTAKR 465  
 QY 483 RVQREKRAVGMGLGAMFLGAGSTMGATSMALTIVQARQLLSGIVQOQNLLRAIQA 542  
 DB 466 RVQREKRAVGMGLGAMFLGAGSTMGATSMALTIVQARQLLSGIVQOQNLLRAIQA 524  
 QY 543 QHLLQITVWGIKQOARILAVERYLKDQQLLGFWGCSGKLICTTAVPWNASNKTLDOI 602  
 DB 543 QHLLQITVWGIKQOARILAVERYLKDQQLLGFWGCSGKLICTTAVPWNASNKTLDOI 602

525 QRMQLQTVMGIKQLQARVLAVERYLQDQLGIWGCCKLICCTAVPWNASWSNKSILDR 584  
 603 WNNMTWMDREIDNTHLYTLIESQKQKQELLQDKWASL 649  
 585 WNNMTWMDREIDNTHLYTLIESQKQKQELLQDKWASL 631  
 RESULT 11  
 AAP80967  
 ID AAP80967 standard; protein; 851 AA.  
 AC AAP80967;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 15-NOV-1990 (first entry)  
 XX  
 DE HIV protein HT6.  
 XX  
 KW HIV; HT6; gp160; envelope protein; RF; AIDS.  
 XX  
 OS Human immunodeficiency virus variant RF.  
 XX  
 PN EP272858-A.  
 PD 29-JUN-1988.  
 XX  
 PF 14-DEC-1987; 87EP-0310967.  
 XX  
 PR 31-AUG-1987; 87US-0091481.  
 PR 15-DEC-1986; 86US-0941111.  
 XX  
 PA (REPK ) REPLIGEN CORP.  
 XX  
 PI Rusche J, Lynn D, Carson H, Putney S, Jellis CL;  
 XX  
 WPI; 1988-176944/26.  
 DR N-PSDB; AAN80948.  
 XX  
 PT Prodn. of recombinant HIV envelope proteins in insect cells -  
 PT useful as vaccine against AIDS and for diagnosis and therapy.  
 XX  
 PS Disclosure; Page ?; 4pp; English.  
 XX  
 CC The sequence is the result of cloning a hybrid envelope gene from  
 CC HIV variants BH10 and RF. A central portion of the RF gene was  
 CC used, the rest being from te distantly related variant BH10. The  
 CC resulting clone, pACHR6 produces a hybrid gp 160 envelope protein  
 CC with novel immunological and antigenic characteristics. It may  
 CC be used to as a vaccine and for diagnosis and therapy of AIDS.  
 CC See also AAP80966.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 851 AA;  
 Query Match 81.1%; Score 2816; DB 9; Length 851;  
 Best Local Similarity 82.0%; Pred. No. 6.9e-151;  
 Matches 533; Conservative 42; Mismatches 53; Indels 22; Gaps 7;  
 3 LWTVYGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62  
 34 LWTVYGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 93  
 63 NNMKNMVEQMEHDIISLWQSLKPCVKLTPLCVLNCITDLNNTNTTELSIIVWVQ 122  
 94 NNMKNMVEQMEHDIISLWQSLKPCVKLTPLCVLNCITDLNNTNTSSSGRIME- 150  
 123 RGKGMNRCNFTTIRDKVOREYALFYKLDVEPIDDDNKNNTTKYELINCNTSVITQ 182  
 151 --XGEIKNCNFMNSTRIRKQVEYAFYKLDIIPIDN-----DTTSTLTSCNTSVITQ 203  
 183 ACPKVSFEPIHYCTPTGTFALLKCNKKFNKGTGCTNVTSTVQCTHGIRPVVSTQLLNG 242  
 204 ACPKVSFEPIHYCAPAGFAILKCNKNTFNGTGTCTNVTSTVQCTHGIRPVVSTQLLNG 263

QY 243 SLAEEVWIRSENFTNAKTIIVQLNVSVVEINCTRNHNHTRKRVTLGPGRVWTTGILG 302  
 DB 264 SLAEEVWIRSENFTNAKTIIVQLNVSVVEINCTRNHNHTRKRVTLGPGRVWTTGILG 323  
 QY 303 NROAHCNISRAQWNTTLOQIATTLREOFQGNKTIAFNQSNGGDPPEIVMHSFNCGBEFYC 362  
 DB 324 DIRAHCNLSRAQWNTTLOQIATTLREOFQGNKTIAFNQSNGGDPPEIVMHSFNCGBEFYC 383  
 QY 363 NSTQLFNSAWNT--SNGTWSVTRKOKTCDITLPCRIKQIINRWQVVGKAMVALPIKG 420  
 DB 384 NTTQLFNSAWNT--SNGTWSVTRKOKTCDITLPCRIKQIINRWQVVGKAMVALPIKG 435  
 QY 421 LIRCSSNITGLLLTRDGGGE-NOTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTK 479  
 DB 436 QIKCISNITGLLLTRDGGEDTNTTEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPTK 495  
 QY 480 AKRRVVOREKRAVGMGLGAMFLGAGSTMGATSMALTVOARQLLSGIVQQNNLLRAI 539  
 DB 496 AKRRVVOREKRAVG-IGALFLGAGSTMGATSMALTVOARQLLSGIVQQNNLLRAI 554  
 QY 540 KAOQHLLQTVMGIKQLQARILAVERYLKDQQLLGFWGCCKLICCTAVPWNASWSNKSIL 599  
 DB 555 EAOQHLLQTVMGIKQLQARILAVERYLKDQQLLGFWGCCKLICCTAVPWNASWSNKSIL 614  
 QY 600 DQIWNMTWMDREIDNTHLYTLIESQKQKQELLQDKWASL 649  
 DB 615 EQIWNMTWMDREIDNTHLYTLIESQKQKQELLQDKWASL 664  
 RESULT 12  
 AAW53112  
 ID AAW53112 standard; Protein; 855 AA.  
 AC AAW53112;  
 XX  
 DT 25-JUN-1998 (first entry)  
 XX  
 DE ENV protein contained in a complete ARV-2 nucleotide sequence.  
 XX  
 KW ARV-2; enhanced promoter; gene expression; cytomegalovirus;  
 XX HIV; AIDS; ENV protein.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN US5688688-A.  
 XX  
 PD 18-NOV-1997.  
 XX  
 PF 10-AUG-1994; 94US-0288336.  
 XX  
 PR 24-DEC-1987; 87US-0138894.  
 PR 31-OCT-1984; 84US-0667501.  
 PR 30-JAN-1985; 85US-0696534.  
 PR 06-SEP-1985; 85US-0773447.  
 PR 17-AUG-1992; 92US-0931191.  
 PR 28-JUN-1993; 93US-0083391.  
 PR 17-AUG-1993; 93US-0107377.  
 PR 10-AUG-1994; 94US-0288336.  
 XX (CHIR ) CHIRON CORP.  
 XX  
 PI Chapman BS, Dina D, Haigwood NL, Luciw PA, Rosenberg S;  
 PI Thayer RM;  
 XX  
 WPI; 1998-007982/01.  
 DR N-PSDB; AAV04733.  
 XX  
 PT Enhanced promoter for gene expression - comprising cytomegalovirus  
 PT immediate early promoter plus intron  
 XX  
 PS Example 1; Fig 4C-P; 9pp; English.  
 XX

CC This sequence represents the ENV protein contained in a complete  
 CC nucleotide sequence of ARV-2 derived from partial sequences of several  
 CC ARV clones. The invention provides a method for construction of a vector  
 CC for expression of a polypeptide in a mammalian cell, comprising a  
 CC polypeptide coding sequence operably linked downstream of an enhanced  
 CC promoter. The enhanced promoter comprises the human cytomegalovirus  
 CC immediate early region (HCMV IE1) promoter and the first intron proximate  
 CC to the 3' end of the HCMV IE1 promoter. The polypeptide can be any of the  
 CC HIV recombinant polypeptides and especially HIV gp120. Expression of HIV  
 CC gp120 by COS 7 cells transfected with pCMV6a containing the gp120 coding  
 CC region, where pCMV6a is a vector containing the above enhanced promoter,  
 CC is increased by a factor of 50-100 compared with the use of a vector  
 CC containing the SV40 early promoter.

XX  
 XX Sequence 855 AA;  
 Query Match 81.0%; Score 2815.5; DB 19; Length 855;  
 Best Local Similarity 81.2%; Pred. No. 7.4e-151;  
 Matches 528; Conservative 47; Mismatches 58; Indels 17; Gaps 6;  
 QY 3 LWTVYGVVPVWKEATTLFCASDAKAYDTAHNVWATHACVPTNPDPQEVVLENTNF 62  
 DB 33 LWTVYGVVPVWKEATTLFCASDAKAYDTAHNVWATHACVPTNPDPQEVVLENTNF 92  
 QY 63 NWWKNNVVEQMHEDIISLWDSQKPCVKLTPLCVTLNCTDLNNTNTTTELSIIIVWVEQ 122  
 DB 93 NWWKNNVVEQMHEDIISLWDSQKPCVKLTPLCVTLNCTDLNNTNTSSN-----WKE 146  
 QY 123 RGKGMRCNSFNITTSIRDKVQREYALFKLDVEPIDDNKNTNTNTKYLINCNTSVITQ 182  
 DB 147 EIKGEIKNCNSFNITTSIRDKIKENALFRNLDDVPIDNASTTNTYVRLIHCRNSVITQ 206  
 QY 183 ACPKVSPFPIPIHYCTPTGPFALLKNDKKFNGTGTCTNVSTVQCTHGRIPVSTQILLNG 242  
 DB 207 ACPKVSPFPIPIHYCTPTGPFALLKNDKKFNGTGTCTNVSTVQCTHGRIPVSTQILLNG 266  
 QY 243 SLAEEVVRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPGRVWYTTGEILG 302  
 DB 267 SLAEEVVRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPGRVWYTTGEILG 326  
 QY 303 NTRQAHCNISRAQWNTLQIATTLREQFG-NKTIAPNQSSGGDPEIIVHGFNCGGEFFY 361  
 DB 327 DIRKAHCNISRAQWNTLQIATTLREQFG-NKTIAPNQSSGGDPEIIVHGFNCGGEFFY 386  
 QY 362 CNSTOLFNSANVVT-SNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKAMYPKIG 420  
 DB 387 CNTTOLFNNWRLNHTGT-----KGNDT---IILPCRIKQIINRWQVVGKAMYPKIG 438  
 QY 421 LIRCSNITGLLLTRDGG-GENQTEIFRPGGDMRDNRSELKYKVVKTEPLGVATPK 479  
 DB 439 QIQCSSNITGLLLTRDGGTNTVTEVFRPGGDMRDNRSELKYKVVKTEPLGVATPK 498  
 QY 480 AKRRVVRQEKRAVGMGLGFLGAAGSTMGATSMALTQVQARQLLSGIVQOQNLLRAI 539  
 DB 499 AKRRVVRQEKRAVGMGLGFLGAAGSTMGATSMALTQVQARQLLSGIVQOQNLLRAI 558  
 QY 540 KAOQHLLQITVWIKIQLOARILAVERYLKDQQLLGFWCGSGKLICTTAVPNWASNKTL 599  
 DB 559 EAOQHLLQITVWIKIQLOARILAVERYLKDQQLLGFWCGSGKLICTTAVPNWASNKSL 618  
 QY 600 DOIWNNTWMEWDREIDNVTIYLLIESQNOQKQOELLQDKWASL 649  
 DB 619 EDIWNNTWQWEREIDNVTIYLLIESQNOQKQOELLQDKWASL 668

RESULT 13

AA77298

ID AA77298 standard; Protein; 855 AA.

XX AA77298;

AC AA77298;

XX 22-MAY-2000 (first entry)

DB HIV-1 (ATCC CRL 8597) env protein.  
 XX HIV-1: immunocassay; antigen; pol fragment; p31; recombinant; antibody;  
 KW detection; env protein.  
 XX Human immunodeficiency virus type 1 'ATCC CRL 8597'.  
 XX US6013432-A.  
 XX 11-JAN-2000.  
 XX 17-MAY-1995; 95US-0443434.  
 PR 08-JUL-1993; 93US-0089407.  
 PR 24-DEC-1987; 87US-0138894.  
 PR 17-AUG-1992; 92US-0931154.  
 PR 31-OCT-1984; 84US-0667501.  
 PR 30-JAN-1985; 85US-0696534.  
 PR 06-SEP-1985; 85US-0773447.  
 XX (CHIR ) CHIRON CORP.  
 PA Luciw PA, Dina D;  
 PI WPI; 2000-170256/15.  
 DR N-PSDB; AAZ90201.  
 XX Immunocassay for antibodies against human immune deficiency virus, for  
 PT diagnosing infection, uses an immunogenic fragment of the pol protein  
 PT as antigen -  
 XX Example 1; Fig 4K-O; 99pp; English.  
 XX The invention relates to the improvement of HIV-1 immunoassays by the  
 CC use of an HIV-1 antigen comprising an immunogenic fragment of  
 CC recombinant or synthetic HIV-1 pol, which is encoded by an approximately  
 CC 9.7 kb sequence between a BstXI restriction site at position 3006 and an  
 CC NdeI site at position 5131 of the genome (the proviral DNA sequence is  
 CC given in AAZ90201). The immunogenic pol fragment is not immunologically  
 CC cross-reactive with human T cell lymphotropic viruses I or II. The  
 CC invention also encompasses the use of p31 as an antigen. The recombinant  
 CC antigens may be produced in Escherichia coli, Saccharomyces cerevisiae or  
 CC in mammalian cells. Immunoassays using the recombinant HIV proteins may  
 CC be used to diagnose and stage HIV-1 infections. Sequences  
 CC AA77294-Y77299 represent proteins encoded by the genome of HIV-1 (ATCC  
 CC CRL 8597).

XX Sequence 855 AA;  
 Query Match 81.0%; Score 2815.5; DB 21; Length 855;  
 Best Local Similarity 81.2%; Pred. No. 7.4e-151;  
 Matches 528; Conservative 47; Mismatches 58; Indels 17; Gaps 6;  
 QY 3 LWTVYGVVPVWKEATTLFCASDAKAYDTAHNVWATHACVPTNPDPQEVVLENTNF 62  
 DB 33 LWTVYGVVPVWKEATTLFCASDAKAYDTAHNVWATHACVPTNPDPQEVVLENTNF 92  
 QY 63 NWWKNNVVEQMHEDIISLWDSQKPCVKLTPLCVTLNCTDLNNTNTTTELSIIIVWVEQ 122  
 DB 93 NWWKNNVVEQMHEDIISLWDSQKPCVKLTPLCVTLNCTDLNNTNTSSN-----WKE 146  
 QY 123 RGKGMRCNSFNITTSIRDKVQREYALFKLDVEPIDDNKNTNTNTKYLINCNTSVITQ 182  
 DB 147 EIKGEIKNCNSFNITTSIRDKIKENALFRNLDDVPIDNASTTNTYVRLIHCRNSVITQ 206  
 QY 183 ACPKVSPFPIPIHYCTPTGPFALLKNDKKFNGTGTCTNVSTVQCTHGRIPVSTQILLNG 242  
 DB 207 ACPKVSPFPIPIHYCTPTGPFALLKNDKKFNGTGTCTNVSTVQCTHGRIPVSTQILLNG 266  
 QY 243 SLAEEVVRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPGRVWYTTGEILG 302  
 DB 267 SLAEEVVRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPGRVWYTTGEILG 326

KW Immunoassay; human immunodeficiency virus; HIV; antibody;  
 KW gag antigen; acquired immunodeficiency syndrome; AIDS; HIV infection;  
 KW vaccine; gene therapy; protein expression system; ARV-2;  
 KW AIDS associated retrovirus 2.

XX OS Retroviridae.  
XX PN US6458527-B1.  
XX PD 01-OCT-2002.  
XX XX  
XX PF 28-JUN-1993; 93US-0083391.  
XX PR 24-DEC-1997; 87US-0138894.  
XX PR 17-AUG-1992; 92US-0931191.  
XX PR 31-OCT-1984; 84US-0667501.  
XX PR 30-JAN-1985; 85US-0696534.  
XX PR 06-SEP-1985; 85US-0773447.  
XX XX  
XX PA (CHIR ) CHIRON CORP.  
XX PI Luciw PA, Dina D, Steimer K, Pescador RS, George-Nascimento C;  
XX PI Parkes D, Hallewell R, Barr PJ, Truett M;  
XX XX  
XX DR WPI; 2003-182063/18.  
XX DR N-ESDB; ABX77296.  
XX XX  
XX PT Immunocassay for detecting the presence of antibodies to a human  
XX PT immunodeficiency virus (HIV) for diagnosing, treating or preventing HIV  
XX PT infection, comprises contacting the sample with an HIV gag antigen -  
XX PS Disclosure; Fig 4K-O; 101pp; English.  
XX XX  
XX CC The invention describes an immunocassay for detecting the presence of  
XX CC antibodies to a human immunodeficiency virus (HIV) in a sample by:  
XX CC contacting the sample with an HIV gag antigen that binds anti-HIV  
XX CC antibodies in an acquired immunodeficiency syndrome (AIDS) patient sera;  
XX CC and determining binding, where the improvement comprises employing as  
XX CC the gag antigen either a synthetic polypeptide or a recombinant  
XX CC polypeptide. The method is used for detecting the presence of antibodies  
XX CC to a human immunodeficiency virus (HIV) in a human sample. The  
XX CC method is used in the preparation of a medicament for diagnosing,  
XX CC treating or preventing HIV infection e.g. a vaccine or for gene  
XX CC therapy. This is the amino acid sequence of a retroviral protein  
XX CC isolated from AIDS associated retrovirus 2 (ARV-2) or HIV.  
XX SQ Sequence 855 AA;  
Query Match 81.0%; Score 2815.5; DB 24; Length 855;  
Best Local Similarity 81.2%; Pred. No. 7.4e-151;  
Matches 528; Conservative 47; Mismatches 59; Indels 17; Gaps 6;  
QY 3 LWVTYYGVVPVWKEATTTLCASDAKAYDTEAHNVWATHACVPTNPDPQEVVLNVTEVF 62  
DB 33 LWVTYYGVVPVWKEATTTLCASDAKAYDTEAHNVWATHACVPTNPDPQEVVLNVTEVF 92  
QY 63 NWKNNMVEQMHEDIISLWQSLKPCVKLTPLCVLTNCTDLNTNTNTTSLIIVVWEQ 122  
DB 93 NWKNNMVEQMHEDIISLWQSLKPCVKLTPLCVLTNCTDLGKATNTSSN-----WKE 146  
QY 123 RKGEMRNCSENIITSIRDKVQREYALPYKLDVDPIDDKNTNTNTKYLINCNSTVITQ 182  
DB 147 EIKGEIKNCSENIITSIRDKVQREYALPYKLDVDPIDDKNTNTNTKYLINCNSTVITQ 206  
QY 183 ACPKVSPEPIPIHCTPTGFAKCKNDKNGTGTCTNVSTVQCTHGRPVVSTQQLLNG 242  
DB 207 ACPKVSPEPIPIHCTPTGFAKCKNDKNGTGTCTNVSTVQCTHGRPVVSTQQLLNG 266  
QY 243 SLABEEVVIRSENFNTNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPGRVWYTTGELG 302  
DB 267 SLABEEVVIRSENFNTNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPGRVWYTTGELG 326  
QY 303 NIRAHCNISPAQWNNTLQIATTLRQFG-NKTIAFNOSGGDPEIVMHSFNCGGEFFY 361  
DB 327 DIRKAHCNISPAQWNNTLQIATTLRQFG-NKTIAFNOSGGDPEIVMHSFNCGGEFFY 386  
QY 362 CNSTQLFNANWNT-SNGTWSVTRKQKDTGDIITLPCRKQIINRWQVWGKMYALPIKG 420

Db 387 CNTTQLENNYWRNLNHTSGT-----KGNDT---IILPCRKQIINRWQVWGKMYALPIKG 438  
QY 421 LIRCSSNITGLLLTRDGG-GENQTEIFRPGGDMRDNWRSELYKYKVVKIEPLGVAPTK 479  
Db 439 QISCSSNITGLLLTRDGGTNTVNDTEVFRPGGDMRDNWRSELYKYKVVKIEPLGVAPTK 498  
QY 480 AKREVVOREKAVGMLGAMFLGFLGAAGSTWGATSMALTVCARQLLSGIVQQNNLLRAI 539  
Db 499 AKREVVOREKAVGMLGAMFLGFLGAAGSTWGATSMALTVCARQLLSGIVQQNNLLRAI 558  
QY 540 KAOQHLLQLTVWGIKQLQARILAVERYLKDQQLLGFWGCCKLICCTTAVPWNASWKNKL 599  
Db 559 EAQOHLIQLTVWGIKQLQARILAVERYLKDQQLLGFWGCCKLICCTTAVPWNASWKNKL 618  
QY 600 DOIWNNTWMEWDREIDNYTHLYTLLEESONQOEKQOEILLQLDKWSL 649  
Db 619 EDIWDNNTWQWEREIDNYTNTIYTLLEESONQOEKQOEILLQLDKWSL 668

Search completed: January 13, 2004, 09:10:01  
Job time : 51 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2004, 09:09:27 ; Search time 22 Seconds  
(without alignments)  
1248.170 Million cell updates/sec

Title: US-09-938-406-1\_COPY\_33\_681

Perfect score: 3474

Sequence: 1 ANLWTVYGVVWKEATT.....QNQKEKNQELLQDKWASL 649

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2852	82.1	850	2	US-08-448-603A-28
2	2852	82.1	850	3	US-09-134-075-28
3	2852	82.1	850	4	US-09-492-739-28
4	2834	81.6	855	1	US-08-022-835-6
5	2834	81.6	855	1	US-08-388-803-6
6	2834	81.6	855	2	US-08-647-714-6
7	2818	81.1	855	3	US-07-956-483-14
8	2814.5	81.0	826	1	US-08-375-510-2
9	2814.5	81.0	826	2	US-08-487-657-2
10	2814.5	81.0	854	4	US-09-309-572-23
11	2814	81.0	880	2	US-08-788-835-7
12	2814	81.0	880	3	US-09-157-963-7
13	2809.5	80.9	856	2	US-07-916-098A-2
14	2809.5	80.9	857	2	US-08-448-603A-30
15	2809.5	80.9	857	3	US-09-134-075-30
16	2809.5	80.9	857	4	US-09-492-739-30
17	2807.5	80.8	855	3	US-07-956-483-15
18	2806.5	80.8	856	3	US-09-124-900-9
19	2806.5	80.8	863	3	US-08-463-210-11
20	2795.5	80.5	856	4	US-09-337-387-11
21	2795	80.5	861	1	US-08-127-499A-14
22	2795	80.5	861	1	US-08-482-847-14
23	2795	80.5	861	3	US-07-956-483-10
24	2795	80.5	861	3	US-08-472-240A-1
25	2795	80.5	861	3	US-08-472-240A-7
26	2795	80.5	861	4	US-08-817-441-103
27	2787	80.2	861	3	US-07-956-483-16

Sequence 13, Appl  
Sequence 5, Appl  
Sequence 32, Appl  
Sequence 32, Appl  
Sequence 32, Appl  
Sequence 2, Appl  
Sequence 10, Appl  
Sequence 6, Appl  
Sequence 3, Appl  
Sequence 12, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 11, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-08-448-603A-28  
; Sequence 28, Application US/08448603A  
; Patent No. 5864027  
; GENERAL INFORMATION:  
; APPLICANT: Berman, Phillip W.  
; APPLICANT: Nakamura, Gerald R.  
; TITLE OF INVENTION: HIV Envelope Polypeptides  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP  
; STREET: 3 Embarcadero Center  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448.603A  
; FILING DATE: 07-JUN-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: 08/072,833  
; FILING DATE: 07-JUN-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halliday, Emily  
; REGISTRATION NUMBER: 38903  
; REFERENCE/DOCKET NUMBER: 14918-704  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-393-2000  
; TELEFAX: 415-393-2286  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 850 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-448-603A-28

Query Match 82.1%; Score 2852; DB 2; Length 850;  
Best Local Similarity 83.0%; Pred. No. 2.4e-227;  
Matches 537; Conservative 42; Mismatches 52; Indels 16; Gaps 7;

3 LWTVYGVVWKEATTTLFCASAKAYDTEAHNVWATHACVPTNPQEVVLENTVF 62

Db 33 LWTVYGVVPWKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQELGLENVTENF 92  
 QY 63 NMKNNNVEQMHEDIISLWQSLKPCVKLPLCVTLNCTDLNNTNTNTTNTTSLIIVVWEQ 122  
 Db 93 NMKNNNVEQMHEDIISLWQSLKPCVKLPLCVTLNCTDL--KNAINTTSS--WGK 146  
 QY 123 RGKEMENCSFNITTSRDKVOREYALFYKLDVPEIDDKNTNTNTKYLRLNCTSVITQ 182  
 Db 147 MERGEIKNCSFNVTISRDKMKNEYALFYKLDVVPIDN-----DNTSYRLISCMNTSVITQ 201  
 QY 183 ACQKVSPEPIHYCTPTGFALLKCNCKFNGTGPCINVSIVQCTHGIRPVVSTQLLING 242  
 Db 202 ACQKVSPEPIHYCAPAGFALLKCRDKKNGTGPCINVSIVQCTHGIRPVVSTQLLING 261  
 QY 243 SLABEEVIRSENPTNNAKTIIVOLNVSVEINCTRPNNHTRKRVTLGPRVWYTTGEILG 302  
 Db 262 SLABEEVIRSEFSDNAKTIIVQLNESVEINCTRPNNNTRRSIHIGGRAFYATGEIIG 321  
 QY 303 NIRAHCNI SPAQNNLTLOIATTLRQFGNKTIAPNOSGGDPEIIVMHSFNCGGEFFYC 362  
 Db 322 DIRQAHNLSSTKWNNTLKQIVTKLREHF-NKTIIVFNHSSGGDPEIIVMHSFNCGGEFFYC 380  
 QY 363 NSTOLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKAMYPALPIKGLI 422  
 Db 381 NITPLFNSWNTY--TWNNTEGSDTGRNITLQCRKQIINMWQEVGKAMYAPPIRQI 438  
 QY 423 RCSSNITGLLLTRDGGGENQTEIFRPGGDMRDNWRSELYKYKVKVIEPIGVAPTAKR 482  
 Db 439 RCSSNITGLLLTRD--GGNNSETEIFRPGGDMRDNWRSELYKYKVKVIEPIGVAPTAKR 497  
 QY 483 RVQREKRAVGLGAMFELGFLGAAGSTMGATSMALTQCAROLLGSIQOQNNLLRAIKAO 542  
 Db 498 RVQREKRAVG-IGAVELGFLGAAGSTMGASVTLTVOARLLLSGIVQOQNNLLRAIEAE 556  
 QY 543 QHLLQLTWGKIQQLQARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASWKNKTLDOI 602  
 Db 557 QHLLQLTWGKIQQLQARILAVERYLKDQQLLGIWCGSGKLICTTAVPWNASWKNKSLDKI 616

## RESULT 2

US-09-134-075-28  
 ; Sequence 28, Application US/09134075  
 ; Patent No. 6042836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berman, Phillip W.  
 ; APPLICANT: Nakamura, Gerald R.  
 ; TITLE OF INVENTION: HIV Envelope Polypeptides  
 ; NUMBER OF SEQUENCES: 33  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP  
 ; STREET: 3 Embarcadero Center  
 ; CITY: San Francisco  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94111  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/134,075  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/448,603  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Haliday, Emily

; REGISTRATION NUMBER: 38903  
 ; REFERENCE/DOCKET NUMBER: 14918-704  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-393-2000  
 ; TELEFAX: 415-393-2286  
 ; TELEX:

; INFORMATION FOR SEQ ID NO: 28:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 850 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-09-134-075-28

Query Match 82.1%; Score 2852; DB 3; Length 850;  
 Best Local Similarity 83.0%; Pred. No. 2.4e-227;  
 Matches 537; Conservative 42; Mismatches 52; Indels 16; Gaps 7;

QY 3 LWTVYGVVPWKEATTTLCASDAKAYDTEAHNVWATHACVPTNPNPQEVLENVTENF 62  
 Db 33 LWTVYGVVPWKEATTTLCASDAKAYDTEVHNWATHACVPTDNPQELGLENVTENF 92  
 QY 63 NMKNNNVEQMHEDIISLWQSLKPCVKLPLCVTLNCTDLNNTNTNTTNTTSLIIVVWEQ 122  
 Db 93 NMKNNNVEQMHEDIISLWQSLKPCVKLPLCVTLNCTDL--KNAINTTSS--WGK 146  
 QY 123 RGKEMENCSFNITTSRDKVOREYALFYKLDVPEIDDKNTNTNTKYLRLNCTSVITQ 182  
 Db 147 MERGEIKNCSFNVTISRDKMKNEYALFYKLDVVPIDN-----DNTSYRLISCMNTSVITQ 201  
 QY 183 ACQKVSPEPIHYCTPTGFALLKCNCKFNGTGPCINVSIVQCTHGIRPVVSTQLLING 242  
 Db 202 ACQKVSPEPIHYCAPAGFALLKCRDKKNGTGPCINVSIVQCTHGIRPVVSTQLLING 261  
 QY 243 SLABEEVIRSENPTNNAKTIIVOLNVSVEINCTRPNNHTRKRVTLGPRVWYTTGEILG 302  
 Db 262 SLABEEVIRSEFSDNAKTIIVQLNESVEINCTRPNNNTRRSIHIGGRAFYATGEIIG 321  
 QY 303 NIRAHCNI SPAQNNLTLOIATTLRQFGNKTIAPNOSGGDPEIIVMHSFNCGGEFFYC 362  
 Db 322 DIRQAHNLSSTKWNNTLKQIVTKLREHF-NKTIIVFNHSSGGDPEIIVMHSFNCGGEFFYC 380  
 QY 363 NSTOLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKAMYPALPIKGLI 422  
 Db 381 NITPLFNSWNTY--TWNNTEGSDTGRNITLQCRKQIINMWQEVGKAMYAPPIRQI 438  
 QY 423 RCSSNITGLLLTRDGGGENQTEIFRPGGDMRDNWRSELYKYKVKVIEPIGVAPTAKR 482  
 Db 439 RCSSNITGLLLTRD--GGNNSETEIFRPGGDMRDNWRSELYKYKVKVIEPIGVAPTAKR 497  
 QY 483 RVQREKRAVGLGAMFELGFLGAAGSTMGATSMALTQCAROLLGSIQOQNNLLRAIKAO 542  
 Db 498 RVQREKRAVG-IGAVELGFLGAAGSTMGASVTLTVOARLLLSGIVQOQNNLLRAIEAE 556  
 QY 543 QHLLQLTWGKIQQLQARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASWKNKTLDOI 602  
 Db 557 QHLLQLTWGKIQQLQARILAVERYLKDQQLLGIWCGSGKLICTTAVPWNASWKNKSLDKI 616  
 QY 603 WNNMTWMEWDREIDNYTHLIIVTLIEESQOQEKQOELLQDLKWSL 649  
 Db 617 WDNMTWMEWEREIDNYTSLIYSLIEESQOQEKNEQELLELDKWSL 663

## RESULT 3

US-09-492-739-28  
 ; Sequence 28, Application US/0942739  
 ; Patent No. 6331404  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berman, Phillip W.  
 ; APPLICANT: Nakamura, Gerald R.  
 ; TITLE OF INVENTION: HIV Envelope Polypeptides  
 ; NUMBER OF SEQUENCES: 33  
 ; CORRESPONDENCE ADDRESS:



ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP  
STREET: 3 Embarcadero Center  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/492,739  
FILING DATE: 27-Jan-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/134,075  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Haliday, Emily  
REGISTRATION NUMBER: 38903  
REFERENCE/DOCKET NUMBER: 14918-704  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-393-2000  
TELEFAX: 415-393-2286  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 850 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 28:

US-09-492-739-28

Query Match 82.1%; Score 2852; DB 4; Length 850;  
Best Local Similarity 83.0%; Pred. No. 2,4e-227;  
Matches 537; Conservative 42; Mismatches 52; Indels 16; Gaps 7;  
QY 3 LWTVTYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPPOEVLENVTENF 62  
DB 33 LWTVTYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPPOEVLENVTENF 92  
QY 63 NWNKNNMVEQHEHDIISLWDSQSLKPCVKLTPLCVTLNCTDLNNTNNTTSLIIVWREQ 122  
DB 93 NWNKNNMVEQHEHDIISLWDSQSLKPCVKLTPLCVTLNCTDLNNTNNTTSSS----WKG 146  
QY 123 RKGEMRNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNNTNNTKYRLNCTNSVITQ 182  
DB 147 MERGEIKNCSEFNITTSIRDKMNEALFYKLDVPIDN-----DNSTYELISCNISVITQ 201  
QY 183 ACPKVSFEPIPIHYCTPTGTFALLKNDKFNCTGCTNVTSTVQCTHGRIPVYSTOLLNG 242  
DB 202 ACPKVSFEPIPIHYCAPAGFALLKCRDKKFNCTGCTNVTSTVQCTHGRIPVYSTOLLNG 261  
QY 243 SLAEVEVIRSENFNTNAKTIIVQLNVSVEINCTPNNHTRKRVTLGPGRVWYTTGEILG 302  
DB 262 SLAEVEVIRSENFNTNAKTIIVQLNVSVEINCTPNNHTRKRVTLGPGRVWYTTGEILG 321  
QY 303 NTRQAHCNISRAQWNNLTQOIAATTLREPGNKTIAPNQSOGDPEIVMHSFNCGGEFFYC 362  
DB 322 DIRQAHCNLSSTKWNLTQIVTKLREHF-NKTIIVFNHSSGGDPEIVMHSFNCGGEFFYC 380  
QY 363 NQTQIFNSAWNTSNGTSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMYPALPIKGLI 422  
DB 381 NTPPLFNSTWNTY--TWNTGSDNTGRNITLQCRKIQIINWQEVGKAMYPALPIRQI 438  
QY 423 RCSSNITGLLLTDCGGENQTTTEIRPGCGDNRDNWRSELYKYKVKIPLGVATTKAKR 482  
DB 439 RCSSNITGLLLTRD--GGNNSETTEIRPGCGDNRDNWRSELYKYKVKIPLGVATTKAKR 497  
QY 483 RVVQREKRAVGMGLGAMFLGFLCAAGSTWGTATSWALTVOARQLLSGIVQOONLLRAIKAO 542

DB 498 RVWQREKRAVG-IGAVFLGFLCAAGSTWGAASVTITVQARLLSGIVQOONLLRAIEAE 556  
QY 543 QHLLQLTWGIKQLQARILAVERYLKDQQLGFWGCGKLICTTAVPMNASWNSKTLDOI 602  
DB 557 QHLLQLTWGIKQLQARILAVERYLKDQQLGFWGCGKLICTTAVPMNASWNSKSLDKI 616  
QY 603 WNNMTWMEWDREIDNYTHLYTLIESQNOQEKNOQELLQDKWASL 649  
DB 617 WDNMTWMEWEIEDNYTSLIYSLIESQNOQEKNEQELLELDKWASL 663  
RESULT 4  
US-08-022-835-6  
Sequence 6, Application US/08022835  
Patent No. 5420030  
GENERAL INFORMATION:  
APPLICANT: Reitz Jr., Marvin S.  
APPLICANT: Franchini, Genoveffa  
APPLICANT: Markham, Phillip D.  
APPLICANT: Gallo, Robert C.  
APPLICANT: Lori, Franco C.  
APPLICANT: Popovic, Mikulas  
APPLICANT: Garnter, Suzanne  
TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1 AND USES  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSER: CUSHMAN, DARBY & CUSHMAN  
STREET: Eleventh Floor, 1615 L. Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036-5601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/022,835  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/599,491  
FILING DATE: 17-OCT-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Scott, Watson T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: WTS/5683/78507/SRL  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 855 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-022-835-6

Query Match 81.6%; Score 2834; DB 1; Length 855;  
Best Local Similarity 82.2%; Pred. No. 7.5e-226;  
Matches 534; Conservative 45; Mismatches 53; Indels 18; Gaps 6;  
QY 3 LWTVTYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPPOEVLENVTENF 62  
DB 33 LWTVTYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPPOEVLENVTENF 92  
QY 63 NWNKNNMVEQHEHDIISLWDSQSLKPCVKLTPLCVTLNCTDLN--TN--NTNTTSLIIV 119  
DB 93 NWNKNNMVEQHEHDIISLWDSQSLKPCVKLTPLCVTLNCTDLNATNGDNTTSSRGW 152

QY 120 WEORGKEMKNCSPNITTSIRDKVOREYALFYKLDVEPIDDKNTNTNTKYRLINCVTSV 179  
 Db 153 -----GGGEMKNCSPNITTNIRKGVQKQYALFYKLDIAPIDNN-----SNNRYRLISCVTSV 204  
 QY 180 ITACPKVSEPIPIHYCTPTGPFALLKCNCKFKNGTGTCTNVSTVQCTHGIRPVVSTQLL 239  
 Db 205 ITACPKVSEPIPIHYCAPAGFAILKCKDKFKNGKGTCTNVSTVQCTHGIRPVVSTQLL 264  
 QY 240 LNSGLAEEVVRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTLGFGRVVYTTGE 299  
 Db 265 LNSGLAEEVVRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTLGFGRVVYTTGE 324  
 QY 300 ILGNIRQAHCHNISRAQWNTLQIATTLREOFNGKTIATFNOSSGDDPEIVHSHSCGGEF 359  
 Db 325 IIGDIRQAHCHNISRAQWNTLQIATTLREOFNGKTIATFNOSSGDDPEIVHSHSCGGEF 384  
 QY 360 FYCNSTQLFNSAMNVTSGTMSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMVAPPIK 419  
 Db 385 FYCNSTQLFNSAMNVTSGTMSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMVAPPIK 438  
 QY 420 GLIRCSSNITGLLTRDGGGENTTEIFRPGGDMRDNWRSELYKYKVVKIEPLGVAPT 479  
 Db 439 GQIRCSNITGLLTRDGGGENTTEIFRPGGDMRDNWRSELYKYKVVKIEPLGVAPT 498  
 QY 480 AKRRVQREKRAVGMGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNNLLRAI 539  
 Db 499 AKRRVQREKRAVG-IGAVFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNNLLRAI 557  
 QY 540 KAOQHLLQLTWGIKQIARLAVERYLKDQQLGFWGCSGKLICTTAVPWNASWNSKTL 649  
 Db 558 EAQOHLQLLTWGIKQIARLAVERYLKDQQLGFWGCSGKLICTTAVPWNASWNSKTL 617  
 QY 600 DQIWNMTWMDREIDNYTHLYTLIEESQOQEKQOQLLDKQWASL 649  
 Db 618 NKIWDNMTWMDREIDNYTHLYTLIEESQOQEKQOQLLDKQWASL 667

## RESULT 5

US-08-388-809-6

; Sequence 6, Application US/08388809

; Patent No. 5576000

; GENERAL INFORMATION:

; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,

; APPLICANT: GENOVEFFA, MARKHAM, PHILIP D. GALLO, ROBERT

; APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND

; APPLICANT: GARTNER, SUZANNE

; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN &amp; FINNEGAN

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK, 3.5"

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/388,809

; FILING DATE: 15-FEB-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/022,835

; FILING DATE: 25-FEB-1993

; APPLICATION NUMBER: US 07/599,491

; FILING DATE: 17-OCT-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: LESLIE A. SERUNIAN

; REGISTRATION NUMBER: 35,353

; REFERENCE/DOCKET NUMBER: 2026-4092US2

## ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 855 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-388-809-6

Query Match 81.6%; Score 2834; DB 1; Length 855;

Best Local Similarity 82.2%; Pred. No. 7.5e-226;

Matches 534; Conservative 45; Mismatches 53; Indels 18; Gaps 6;

QY 3 LWTVVYGVVPVWKEATITLFCASDRKAYDTEAHNVWATHACVPTDPNPQEVLEKNVTENF 62

Db 33 LWTVVYGVVPVWKEATITLFCASDRKAYDTEAHNVWATHACVPTDPNPQEVLEKNVTENF 92

QY 63 NMKNNNVQWHEHDIISLWQSLKPCVKLTPLCVTLNCTDLN--TN-NTNTTSLIIV 119

Db 93 NMKNNNVQWHEHDIISLWQSLKPCVKLTPLCVTLNCTDLNATNGDNTTSSRGV 152

QY 120 WEORGKEMKNCSPNITTSIRDKVOREYALFYKLDVEPIDDKNTNTNTKYRLINCVTSV 179

Db 153 -----GGGEMKNCSPNITTNIRKGVQKQYALFYKLDIAPIDNN-----SNNRYRLISCVTSV 204

QY 180 ITACPKVSEPIPIHYCTPTGPFALLKCNCKFKNGTGTCTNVSTVQCTHGIRPVVSTQLL 239

Db 205 ITACPKVSEPIPIHYCAPAGFAILKCKDKFKNGKGTCTNVSTVQCTHGIRPVVSTQLL 264

QY 240 LNSGLAEEVVRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTLGFGRVVYTTGE 299

Db 265 LNSGLAEEVVRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTLGFGRVVYTTGE 324

QY 300 ILGNIRQAHCHNISRAQWNTLQIATTLREOFNGKTIATFNOSSGDDPEIVHSHSCGGEF 359

Db 325 IIGDIRQAHCHNISRAQWNTLQIATTLREOFNGKTIATFNOSSGDDPEIVHSHSCGGEF 384

QY 360 FYCNSTQLFNSAMNVTSGTMSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMVAPPIK 419

Db 385 FYCNSTQLFNSAMNVTSGTMSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMVAPPIK 438

QY 420 GLIRCSSNITGLLTRDGGGENTTEIFRPGGDMRDNWRSELYKYKVVKIEPLGVAPT 479

Db 439 GQIRCSNITGLLTRDGGGENTTEIFRPGGDMRDNWRSELYKYKVVKIEPLGVAPT 498

QY 480 AKRRVQREKRAVGMGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNNLLRAI 539

Db 499 AKRRVQREKRAVG-IGAVFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNNLLRAI 557

QY 540 KAOQHLLQLTWGIKQIARLAVERYLKDQQLGFWGCSGKLICTTAVPWNASWNSKTL 599

Db 558 EAQOHLQLLTWGIKQIARLAVERYLKDQQLGFWGCSGKLICTTAVPWNASWNSKTL 617

QY 600 DQIWNMTWMDREIDNYTHLYTLIEESQOQEKQOQLLDKQWASL 649

Db 618 NKIWDNMTWMDREIDNYTHLYTLIEESQOQEKQOQLLDKQWASL 667

## RESULT 6

US-08-647-714-6

; Sequence 6, Application US/08647714

; Patent No. 5869313

; GENERAL INFORMATION:

; APPLICANT: REITZ, JR., MARVIN S.; FRANCHINI,

; APPLICANT: GENOVEFFA, MARKHAM, PHILIP D. GALLO, ROBERT

; APPLICANT: C.; LORI, FRANCO C.; POPOVIC, MIKULAS; AND

; APPLICANT: GARTNER, SUZANNE

; TITLE OF INVENTION: MOLECULAR CLONES OF HIV-1

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MORGAN & FINNEGAN  
 STREET: 345 PARK AVENUE  
 CITY: NEW YORK  
 STATE: NEW YORK  
 COUNTRY: USA  
 ZIP: 10154  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY DISK, 3.5"  
 COMPUTER: IBM PC COMPATIBLE  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WORDPERFECT 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/647,714  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/388,809  
 FILING DATE: 15-FEB-1995  
 APPLICATION NUMBER: US/08/022,835  
 FILING DATE: 25-FEB-1993  
 APPLICATION NUMBER: US/07/599,491  
 FILING DATE: 17-OCT-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: LESLIE A. SERUNIAN  
 REGISTRATION NUMBER: 35,353  
 REFERENCE/DOCKET NUMBER: 2026-4092US2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 758-4800  
 TELEFAX: (212) 751-6849  
 TELEX: 421792  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 855 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-647-714-6

Query Match 81.6%; Score 2834; DB 2; Length 855;  
 Best Local Similarity 82.2%; Pred. No. 7.5e-226;  
 Matches 534; Conservative 45; Mismatches 53; Indels 18; Gaps 6;  
 QY 3 LWVTYVGVVWKEATTTLCFASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62  
 DB 33 LWVTYVGVVWKEATTTLCFASDRKAYDTEVHNWATHACVPTNPQEVVLENTNF 92  
 QY 63 NWKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLN--TN--NTNTTTELSIIV 119  
 DB 93 NWKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNATNGNDTNTSSRGWV 152  
 QY 120 WEQKQKGMNRCSPNITTSIRKQVREYALFYKLDVEPIDDKNTNTNTKYRLINCNTSV 179  
 DB 153 ----GGGEMKNGSFNITNIRGVQKAYFYKLDIAPIDNN---SNRRYRLISCNTSV 204  
 QY 180 ITQACPVSFEPIPHYCTPTCFALLKNDKFGNGTCTNVTQCTGIRPVVSTQLL 239  
 DB 205 ITQACPVSFEPIPHYCAPAGFALLKNDKFGNGTCTNVTQCTGIRPVVSTQLL 264  
 QY 240 LNSLAEEVWIRSENFNTNNAKTIIVQLNVSEINCTRNHNTRKRVLTGPGRVVYTCGE 299  
 DB 265 LNSLAEEVWIRSENFNTNNAKTIIVQLNVSEINCTRNHNTRKRVLTGPGRVVYTCGE 324  
 QY 300 ILGNIRQAHCNTRSAQWNTLQOIAATLREQGKNTIAFNQSGGDEPEIVMHSFNCGGEF 359  
 DB 325 ILGDIRQAHCNLSRAKNDTLNKIVIKLREQGKNTIVFKHSSGGDEPEIVMHSFNCGGEF 384  
 QY 360 FYCNSQTQFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMVALPIK 419  
 DB 385 FYCNSQTQFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMVALPIK 438  
 QY 420 GLIRCSSNITGLLTRDGGENOTTEIPRPGGDMRDNRSELYKYKVKVKEPLGVAPTK 479  
 DB 439 GLIRCSSNITGLLTRDGGENOTTEIPRPGGDMRDNRSELYKYKVKVKEPLGVAPTK 498

QY 480 AKRRVVQREKRAVGMGAMFLGFLGAAGSTMGTATMALTVAQRLLSGIVQQNNLLRAI 539  
 DB 499 AKRRVVQREKRAVG-IGAVFLGFLGAAGSTMGTAAAMTLTVAQRLLSGIVQQNNLLRAI 557  
 QY 540 KAAQHLLQTLVWGIKOLOARILAVERYLKDQOLGFWGCGKLICTTAVPWNASWNSKIL 599  
 DB 558 EAQOHLLOTLVWGIKOLOARILAVERYLKDQOLGFWGCGKLICTTAVPWNASWNSKIL 617  
 QY 600 DQIWNMTWMDREIDNTHLYIYIEESQKQKNOQELQLDKWASL 649  
 DB 618 NKIWNMTWMDREIDNTHLYIYIEESQKQKNOQELQLDKWASL 667

RESULT 7  
 US-07-956-483-14  
 ; Sequence 14, Application US/07956483  
 ; Patent No. 6261799  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KIENY, Marie-Paule  
 ; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED  
 ; TITLE OF INVENTION: gp160 VARIANT  
 ; NUMBER OF SEQUENCES: 26  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Burns, Doane, Swecker & Mathis  
 ; STREET: P.O. Box 1404  
 ; CITY: Alexandria  
 ; STATE: Virginia  
 ; COUNTRY: United States  
 ; ZIP: 22313-1404  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/956,483  
 ; FILING DATE: 31-DEC-1992  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: WO 92/19742  
 ; FILING DATE: 12-NOV-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: FR 91 05392  
 ; FILING DATE: 02-MAY-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Crane-Feury, Sharon E  
 ; REGISTRATION NUMBER: 36,113  
 ; REFERENCE/DOCKET NUMBER: 017753-005  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 836-6620  
 ; TELEFAX: (703) 836-2021  
 ; INFORMATION FOR SEQ ID NO: 14:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 855 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; US-07-956-483-14

Query Match 81.1%; Score 2818; DB 3; Length 855;  
 Best Local Similarity 81.3%; Pred. No. 1.6e-224;  
 Matches 534; Conservative 39; Mismatches 54; Indels 30; Gaps 7;  
 QY 3 LWVTYVGVVWKEATTTLCFASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62  
 DB 33 LWVTYVGVVWKEATTTLCFASDAKAYDTEVHNWATHACVPTNPQEVVLENTNF 92  
 QY 63 NWKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVWVEQ 122  
 DB 93 NWKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDLNNTNTTSTNATNTS----SN 148

```

QY 123 RGK---GEMRNCSENIITSIRDKVOREYALFYKLDVDPIDDKNTTNTTKYRLINCNTSV 179
Db 149 RGRVGGEMTNCSENIITSIRSKVQKEYALFYKLDVVPID-----NTSTLINCNTSV 201
QY 180 ITQACPVSFPIPIHYCTPTGFALLKCNDRKFKNGTGCTNVSTVQCTHGIRPVVSTQLL 239
Db 202 ITQACPVSFPIPIHYCA-RWFALLCNCKFKNGTGCTNVSTVQCTHGIRPVVSTHLL 260
QY 240 LNSLABEEVVRSENFNTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYITGE 299
Db 261 LNSLABEEVVRSENFNTNNAKTIIVQLKEAVEINCTRPNNHTRSIHIGPGRFYATGD 320
QY 300 ILGNIRQAHCHNISRAQWNTTLOQIATTLREOPGNKTIAFNOSGSDPPIVMSHSCGGEF 359
Db 321 IIGDIRQAHCHNISRAKWNATLQIVIKURDQFENKTIIFNRSOGSDPPIVMSHSCGGEF 380
QY 360 FYCNSTOLFNSAWNT--SNGTWSVTRKQKDTGDIITLPCRICKQIINRWQVVGKAMYALP 417
Db 381 FYCNSTOLFNSAWNTGG-----NDTITLPCRICKQIINRWQVVGKAMYAPP 432
QY 418 IKGLIRCSNITGLLLTRDQ-----GGNQTEIFRPGGDMRDNRSELYKYKVVKIEP 472
Db 433 IKGVKCSNITGLLLTRDQSGNKGSKNENETIFRPGGDMRDNRSELYKYKVVKIEP 492
QY 473 LGVAPTAKRRVQREKRAVGLGAMFLGELGAAGSTMGATSMALTQVQARQLLSGIVQQQ 532
Db 493 LGVAPTAKRRVQREKRAVGTIGAMFLGELGAAGSTMGATSMALTQVQARQLLSGIVQQQ 552
QY 533 NNILRAIKAQOHLQLTWGIGKQLOARILAVERYLKDOQLLGFWGCCKGLICTTAVPWNNA 592
Db 553 NNILRAIEAQOHLQLTWGIGKQLOARILAVERYLKDOQLLGFWGCCKGLICTTAVPWN 612
QY 593 SWSNKTLDQIWNMTWMEWDREIDNYTHLYTLIESONQOEKQOELLQLDKWSL 649
Db 613 SWSNKSILDKIWNMTWMEWDREIDNYTHLYTLIESONQOEKQOELLQLDKWSL 669

```

## RESULT 8

```

US-08-375-510-2
; Sequence 2, Application US/08375510
; Patent No. 5576421
; GENERAL INFORMATION:
; APPLICANT: Saito, Atsushi
; APPLICANT: Sinagawa, Hideo
; APPLICANT: Nakata, Atsuo
; TITLE OF INVENTION: HIV ANTIGEN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/375,510
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/985,949
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Svensson, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 216-309P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848

```

```

; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 826 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-375-510-2

Query Match      81.0%; Score 2814.5; DB 1; Length 826;
Best Local Similarity 82.9%; Pred. No. 2.9e-224;
Matches 540; Conservative 34; Mismatches 56; Indels 21; Gaps 9;

QY 3 LWTVYYGVVPWKATTTLFCASDAKAYDTAEHNVWATHACVTPNPDPQEVVLNVTFNF 62
Db 6 LWTVYYGVVPWKATTTLFCASDAKAYDTAEHNVWATHACVTPNPDPQEVVLNVTFNF 65
QY 63 NMKNNNVQEQHEDIISLDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVVMEQ 122
Db 66 NMKNNDVVEQHEDIISLDQSLKPCVKLTPLCVSLKCTDL--KNDNTNSSSGRMINE- 122
QY 123 RKGEMRNCSENIITSIRDKVOREYALFYKLDVDPIDDKNTTNTTKYRLINCNTSVITQ 182
Db 123 --KGEIKNCSENIITSIRDKVQKEYAFYKLDIVPID-----NTSYRLISCNTSVITQ 173
QY 183 ACPKVSFPIPIHYCTPTGFALLKCNDRKFKNGTGCTNVSTVQCTHGIRPVVSTQLLNG 242
Db 174 ACPKVSFPIPIHYCAPAGFAILKCNKNTFNGTGCTNVSTVQCTHGIRPVVSTQLLNG 233
QY 243 SLAEDEVVIRSENTNNAKTIIVQLNVSEINCTRPNNHTEK--RVTLGPRVWYITGEI 300
Db 234 SLAEDEVVIRSENFNTNNAKTIIVQLNVTSEINCTRPNNHTEKRIQSGPRAFVITGKI 293
QY 301 LGNIRQAHCHNISRAQWNTTLOQIATTLREQFG-NKTIAFNOSGSDPPIVMSHSCGGEF 359
Db 294 -GNMRQAHCHNISRAKWNATLQIASKUREQEGNNKTIIFKQSSGGDPPIVTHSFNCGGEF 352
QY 360 FYCNSTOLFNSAWNTSNGTWSVTRKQKDTG-DIITLPCRICKQIINRWQVVGKAMYALPI 418
Db 353 FYCNSTOLFNSW--FNSTWSTEGSNNTGSDTITLPCRICKQFINMWQEVGKAMYAPPI 409
QY 419 KGLIRCSNITGLLLTRDQGGENQTTETFRPGGDMRDNRSELYKYKVVKIEPLGVAPT 478
Db 410 SGQIRCSNITGLLLTRDQGGNNNGSEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPT 469
QY 479 KAKRRVQREKRAVGLGAMFLGELGAAGSTMGATSMALTQVQARQLLSGIVQQNNLLRA 538
Db 470 KAKRRVQREKRAVG-IGALFLGLGAAGSTMGCTSMALTQVQARQLLSGIVQQNNLLRA 528
QY 539 IKAQOHLQLTWGIGKQLOARILAVERYLKDOQLLGFWGCCKGLICTTAVPWNASWSNKT 598
Db 529 IEAQOHLQLTWGIGKQLOARILAVERYLKDOQLLGFWGCCKGLICTTAVPWNASWSNKS 588
QY 599 LDOJWNMTWMEWDREIDNYTHLYTLIESONQOEKQOELLQLDKWSL 649
Db 589 LEQIWNMTWMEWDREIDNYTHLSLIESONQOEKQOELLQLDKWSL 639

```

## RESULT 9

```

US-08-487-657-2
; Sequence 2, Application US/08487657
; Patent No. 5834267
; GENERAL INFORMATION:
; APPLICANT: Saito, Atsushi
; APPLICANT: Sinagawa, Hideo
; APPLICANT: Nakata, Atsuo
; TITLE OF INVENTION: HIV ANTIGEN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch and Birch

```

STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,657  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/375,510  
FILING DATE: 18-JAN-1995  
APPLICATION NUMBER: US/07/985,949  
FILING DATE: 04-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 216-309P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 241-1300  
TELEFAX: (703) 241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 826 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
US-08-487-657-2

Query Match 81.0%; Score 2814.5; DB 2; Length 826;  
Best Local Similarity 82.9%; Pred. No. 2.9e-224;  
Matches 540; Conservative 34; Mismatches 56; Indels 21; Gaps 9;  
QY 3 LWTVYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPQVWLVNVTENF 62  
DB 6 LWTVYGVVWKEATTLFCASDAKAYDEVHNVWATHACVPTNPQVWLVNVTENF 65  
QY 63 NMKNMNVQMHEDIISLWDSKPCVKLTPLCVLNCNTLNNTNTNTTSLIIVVVEQ 122  
DB 66 NMKNMNVQMHEDIISLWDSKPCVKLTPLCVLNCNTLNNTNTNTTSLIIVVVEQ 122  
QY 123 RGKGMNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCVTSVITQ 182  
DB 123 --KGEIKNCSFNITTSIRDKVQREYALFYKLDVPIID--NTSYRLISCVTSVITQ 173  
QY 183 ACPKVSFEPIPIHYCTPTGFPALLKCNDFKNGTGPCTNVSTVQCTHGIRPVSTQLLNG 242  
DB 174 ACPKVSFEPIPIHYCAPAGFAILLKCNKNTFNGTGPCTNVSTVQCTHGIRPVSTQLLNG 233  
QY 243 SLAEEVWIRSENFTNAKTIIVQLNVSVENCTRNPNHTRK--RVTLGPRVWYTTGEI 300  
DB 234 SLAEEVWIRSANFTDNAKTIIVQLNVSVEINCTRNPNHTRK--RVTLGPRVWYTTGEI 293  
QY 301 LGNIROAHNISRAQWNTLQIATTLREQFG--NKTIAFNQSSGGDPEIVMHSFNCGGEF 359  
DB 294 --GNMROAHNISRAKWNTLQIASKLREQFGNKTIIIFKQSSGGDPEIVTHSFNCGGEF 352  
QY 360 FYCNSSTGLTLLTRDGGNGTTFIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 418  
DB 353 FYCNSSTGLFNSW---FNSWSTEGSNTEGSDTITLPCRIKQFINWQEVGKAMYAPI 409  
QY 419 KGLIRCSNITGLLLTRDGGNGTTFIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 478  
DB 410 SGQIRCSNITGLLLTRDGGNNNGSEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 469

QY 479 KAKRRVQREKRAVGMGLGAMFLGFIAGAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538  
DB 470 KAKRRVQREKRAVG--IGALFLGFIAGAGSTMGCTSMTLTVOARQLLSGIVQOQNLLRA 528  
QY 539 IKAQOHLQLTWGIGIKOLQARILAVERYLKDQQLGFWGCSGKLICTTAVPWNASWSNKT 598  
DB 529 IEAQOHLQLTWGIGIKOLQARILAVERYLKDQQLGFWGCSGKLICTTAVPWNASWSNKS 588  
QY 599 LDQIWNMTWMEWDREIDNTHLYTLIESQOQEKQOELLQLDKWSL 649  
DB 589 LEQIWNMTWMEWDREINNYTSLIHSLEESQOQEKQOELLQLDKWSL 639  
RESULT 10  
US-09-309-572-23  
Sequence 23, Application US/09309572  
Patent No. 6440730  
GENERAL INFORMATION:  
APPLICANT: Heinrich-Pette-Institut  
TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV  
FILE REFERENCE: P50489  
CURRENT APPLICATION NUMBER: US/09/309,572  
CURRENT FILING DATE: 1999-05-11  
EARLIER APPLICATION NUMBER: DE 198 56 463  
EARLIER FILING DATE: 1998-11-26  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 23  
LENGTH: 854  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
FEATURE:  
OTHER INFORMATION: envelope polyprotein  
US-09-309-572-23

Query Match 81.0%; Score 2814.5; DB 4; Length 854;  
Best Local Similarity 82.9%; Pred. No. 3.1e-224;  
Matches 540; Conservative 34; Mismatches 56; Indels 21; Gaps 9;  
QY 3 LWTVYGVVWKEATTLFCASDAKAYDEAHNVWATHACVPTNPQVWLVNVTENF 62  
DB 34 LWTVYGVVWKEATTLFCASDAKAYDEVHNVWATHACVPTNPQVWLVNVTENF 93  
QY 63 NMKNMNVQMHEDIISLWDSKPCVKLTPLCVLNCNTLNNTNTTSLIIVVVEQ 122  
DB 94 NMKNMNVQMHEDIISLWDSKPCVKLTPLCVLNCNTLNNTNTNTSSSGRIME- 150  
QY 123 RGKGMNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCVTSVITQ 182  
DB 151 --KGEIKNCSFNITTSIRDKVQREYALFYKLDVPIID--NTSYRLISCVTSVITQ 201  
QY 183 ACPKVSFEPIPIHYCTPTGFPALLKCNDFKNGTGPCTNVSTVQCTHGIRPVSTQLLNG 242  
DB 202 ACPKVSFEPIPIHYCAPAGFAILLKCNKNTFNGTGPCTNVSTVQCTHGIRPVSTQLLNG 261  
QY 243 SLAEEVWIRSENFTNAKTIIVQLNVSVENCTRNPNHTRK--RVTLGPRVWYTTGEI 300  
DB 262 SLAEEVWIRSANFTDNAKTIIVQLNVSVEINCTRNPNHTRK--RVTLGPRVWYTTGEI 321  
QY 301 LGNIROAHNISRAQWNTLQIATTLREQFG--NKTIAFNQSSGGDPEIVMHSFNCGGEF 359  
DB 322 --GNMROAHNISRAKWNTLQIASKLREQFGNKTIIIFKQSSGGDPEIVTHSFNCGGEF 380  
QY 360 FYCNSSTGLTLLTRDGGNGTTFIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 418  
DB 381 FYCNSSTGLFNSW---FNSWSTEGSNTEGSDTITLPCRIKQFINWQEVGKAMYAPI 437  
QY 419 KGLIRCSNITGLLLTRDGGNGTTFIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 478  
DB 438 SGQIRCSNITGLLLTRDGGNNNGSEIFRPGGDMRDNRSELYKYKVKVIEPLGVAPT 497  
QY 479 KAKRRVQREKRAVGMGLGAMFLGFIAGAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538

Db 498 KAKRRVQEKRAVG-IGALFLGFLGAAGSTMGCTSMITVQARQLSDIVQOQNLLRA 556  
QY 539 IKAQOHLQLTVWGKIQLOARILAVERYLKQOQLLGFWCGSGKLICTTAVPWNASWNTK 598  
Db 557 IEAQOHLQLTVWGKIQLOARILAVERYLKQOQLLGFWCGSGKLICTTAVPWNASWNTK 616  
QY 599 LQIWNMTWMDREIDNYTHLYITLIESQOQKQOQLLQDLKWSL 649  
Db 617 LEQIWNMTWMDREIDNYTHLYITLIESQOQKQOQLLQDLKWSL 667

RESULT 11  
US-08-788-815-7  
; Sequence 7, Application US/08788815  
; Patent No. 5846546  
; GENERAL INFORMATION:  
; APPLICANT: St. Jude Children's Research Hospital  
; APPLICANT: 332 No. 584546th Lauderdale  
; APPLICANT: PO Box 318  
; APPLICANT: Memphis, TN 38101-0318  
; APPLICANT: United States of America  
; APPLICANT: Coleclough, Christopher  
; APPLICANT: Owens, Randall J.  
; APPLICANT: Slobod, Karen  
; TITLE OF INVENTION: PREPARATION AND USE OF VIRAL VECTORS FOR  
; TITLE OF INVENTION: MIXED ENVELOPE PROTEIN VACCINES AGAINST HUMAN  
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: KLAUBER & JACKSON  
; STREET: 411 HACKENSACK AVENUE  
; CITY: HACKENSACK  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/788,815  
; FILING DATE: Herewith  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul F. Fehlner  
; REGISTRATION NUMBER: 35,135  
; REFERENCE/DOCKET NUMBER: 1340-1-011CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5500  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 880 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: protein  
US-08-788-815-7

Query Match 81.0%; Score 2814; DB 2; Length 880;  
Best Local Similarity 82.1%; Pred. No. 3.5e-224;  
Matches 540; Conservative 39; Mismatches 61; Indels 18; Gaps 9;  
QY 3 LWTVYGVVPVWKEATITLFCASDAKAYDTEAHNVWATHACVTPNPNQEVVLENTVF 62  
Db 43 LWTVYGVVPVWKEATITLFCASDAKAYDTEAHNVWATHACVTPNPNQEVVLENTVF 102  
QY 63 NMWKNVVEQMHEDIISLWDSQKPCVKLTPLCVLNTCTDL--NNNTNTTETLSIIWV 120  
Db 103 NMWKNVVEQMHEDIISLWDSQKPCVKLTPLCVLNTCTDLKNDTNTNNVTSSS---W 158

QY 121 EQR--CKGEMRNCSEFNITTSIRDKVQREYALFYKLDVEPID--DNKNTNTNTKYLINCN 176  
Db 159 GRNMEEGEIKNCSEFNITTSIRDKVQREYALFYKLDIIPIDKNDSDNTTSYKFTITSCN 218  
QY 177 TSVITQACPKVSEFPIPIHYCTPTGALLKNDKFKNGTGPCTNVTSTVQCTHGIRPVVST 236  
Db 219 TSVITQACPKVSEFPIPIHYCAPAGFAILKCNKTFNGTGPCTNVTSTVQCTHGIRPVVST 278  
QY 237 QLLNGSLAEERVWIRSENFNTNAKTIIVQLNVSEINCTRENNTNRK--RVTLGSGRVW 294  
Db 279 QLLNGSLAEERVWIRSENFNTNAKTIIVQLNQSEINCTRENNTNRKIRIQRGFRAP 338  
QY 295 YTTGILGNIRQAHNCISRAOWNNTLQIATTLREQFG-NKTIAPNOSGGDPEIWMHSF 353  
Db 339 VTIGKILGNWRQAHNCISRAOWNNTLQIDSKLREQFGNKKIIFKQSSGGDPEIWMHSF 398  
QY 354 NCGGEFFYCNSTQLFNSAWNVTSTGVTWVTRKQKDTG-DIITLPCRIKQIINRWQVVGKA 412  
Db 399 NCGGEFFYCNSTQLFNSAWNVTSTGVTWVTRKQKDTG-DIITLPCRIKQIINRWQVVGKA 455  
QY 413 MYALPIKGLIRCSSNITGLLTRDGG-GENQTTETIFRPGGDMRDNWSELYKYKVVKIE 471  
Db 456 MYALPIKGLIRCSSNITGLLTRDGG-GENQTTETIFRPGGDMRDNWSELYKYKVVKIE 515  
QY 472 PLGVAPTKAKRVVQREKRAVGMFLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQ 531  
Db 516 PLGVAPTKAKRVVQREKRAVGMFLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQ 575  
QY 532 QNLLRAIKAOQHLLQLTVWGKIQLOARILAVERYLKQOQLLGFWCGSGKLICTTAVPWN 591  
Db 576 QNLLRAIKAOQHLLQLTVWGKIQLOARILAVERYLKQOQLLGFWCGSGKLICTTAVPWN 635  
QY 592 ASWSNKTLDQIWNMTWMDREIDNYTHLYITLIESQOQKQOQLLQDLKWSL 649  
Db 636 ASWSNKTLDQIWNMTWMDREIDNYTHLYITLIESQOQKQOQLLQDLKWSL 693

RESULT 12  
US-09-157-963-7  
; Sequence 7, Application US/09157963  
; Patent No. 6086891  
; GENERAL INFORMATION:  
; APPLICANT: St. Jude Children's Research Hospital  
; APPLICANT: 332 No. 6086891th Lauderdale  
; APPLICANT: PO Box 318  
; APPLICANT: Memphis, TN 38101-0318  
; APPLICANT: United States of America  
; APPLICANT: Coleclough, Christopher  
; APPLICANT: Owens, Randall J.  
; APPLICANT: Slobod, Karen  
; TITLE OF INVENTION: PREPARATION AND USE OF VIRAL VECTORS FOR  
; TITLE OF INVENTION: MIXED ENVELOPE PROTEIN VACCINES AGAINST HUMAN  
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: KLAUBER & JACKSON  
; STREET: 411 HACKENSACK AVENUE  
; CITY: HACKENSACK  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/157,963  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Paul F. Fehlner  
; REGISTRATION NUMBER: 35,135  
; REFERENCE/DOCKET NUMBER: 1340-1-011CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5500  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 880 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: protein  
US-09-157-963-7



QY 419 KGLRCSNITGLLLTRDGGGNOTTEIFRPGGDMRDNWRSELYKYYKWKIEPLGVAPT 478  
DB 440 SQQRCSNITGLLLTRDGGGNSNESEIFRPGGDMRDNWRSELYKYYKWKIEPLGVAPT 499  
QY 479 KAKRRVQREKRAVGMGLGAMFLGFGAAGSTMGATSMALTVOAROLLGSIQQNNLLRA 538  
DB 500 KAKRRVQREKRAVG-IGALFLGFLGAAGSTMGAASMTLTVOAROLLGSIQQNNLLRA 558  
QY 539 IKAQOHLIQLTQWGIKQIARILAVERYLKQOQLLGFPGCSGKLICTTAVPNWASWSNKT 598  
DB 559 IEAQOHLIQLTQWGIKQIARILAVERYLKQOQLLGFPGCSGKLICTTAVPNWASWSNKS 618  
QY 599 LDOIWNMTWMEWDEIDNTHLYTLIEESQOQKNOQELLDKQWASL 649  
DB 619 LEQIWNHTTWMEWDEIDNTHLYTLIEESQOQKNOQELLDKQWASL 669

RESULT 14  
US-08-448-603A-30  
; Sequence 30, Application US/08448603A  
; Patent No. 5864027  
; GENERAL INFORMATION:  
; APPLICANT: Berman, Phillip W.  
; APPLICANT: Nakamura, Gerald R.  
; TITLE OF INVENTION: HIV Envelope Polypeptides  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP  
; STREET: 3 Embarcadero Center  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fast-SEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/448,603A  
; FILING DATE: 07-JUN-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/072,833  
; FILING DATE: 07-JUN-93  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halliday, Emily  
; REGISTRATION NUMBER: 38903  
; REFERENCE/DOCKET NUMBER: 14918-704  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-393-2000  
; TELEFAX: 415-393-2286  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 857 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-448-603A-30

Query Match 80.9%; Score 2809.5; DB 2; Length 857;  
Best Local Similarity 81.3%; Pred No. 7.9e-224;  
Matches 529; Conservative 45; Mismatches 60; Indels 17; Gaps 6;

QY 3 LWVTYYGVPMKEATTTLCFASDAXDYDEAHNWMATHACVPTNPNFQEVVLENTVF 62  
DB 33 LWVTYYGVPMKETTTLTLCFASDAXDYDETHNWMATHACVPTDPNPFQEVVLENTVF 92  
QY 63 NWNKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELIIIVNEQ 122  
DB 93 NWNKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD--AGNTTNTNSSS----REK 146

QY 123 RKGEMRNCSENIITTSIRDKVQREYALFYKLDVEFIDN--KNTTNTKYILINCMTSVI 180  
DB 147 LEKGEIKNCSENIITTSVEDKVKQKETAFLFNKLDIVFDODDRNSTRNSTNYRLISCNVSI 206  
QY 181 TOACPQVSEFPIPIHYCTPTGTFALLKNDKXFNKGTCTNVTSTVQCTGIRPVVSTQLLL 240  
DB 207 TOACPQVSEFPIPIHFCTPAGFALLKCNKNTFNKGTCTNVTSTVQCTGIRPVVSTQLLL 266  
QY 241 NGSLAEERWVIRSENFNTNNAKTIIVQLNVSVEINCTRENNTNKRVTIPGRCVWVYTTGEI 300  
DB 267 NGSLAEERWVIRSENFNTNNAKTIIVQLTEPKINCTRENNTNKRKIPGRCVAFATGDI 326  
QY 301 LGNIRQAHCNISRAQWNNLTQOIAITLREQFNGKNTIAFNQSSGGDPEIWMHSFNCGGEFF 360  
DB 327 LGNIRQAHCNLSRTDNNNTLQIIVKELREQFNGKNTIIEHSSGGDPEIWMHSFNCRGEFF 386  
QY 361 YCNSTQLPNSAWNVT--SNGTWSVTRKOKDTGDIITLPCRKIKOINRQVVGKAMVALPI 418  
DB 387 YCNSTQLPNSAWNVT--SNGTWSVTRKOKDTGDIITLPCRKIKOINRQVVGKAMVALPI 440  
QY 419 KGLRCSNITGLLLTRDGGGNOTTEIFRPGGDMRDNWRSELYKYYKWKIEPLGVAPT 478  
DB 441 SQQRCSNITGLLLTRDGGGNSNESEIFRPGGDMRDNWRSELYKYYKWKIEPLGVAPT 500  
QY 479 KAKRRVQREKRAVGMGLGAMFLGFGAAGSTMGATSMALTVOAROLLGSIQQNNLLRA 538  
DB 501 KAKRRVQREKRAVG-IGAVFLGFLGAAGSTMGAASITLTVOAROLLGSIQQNNLLRA 559  
QY 539 IKAQOHLIQLTQWGIKQIARILAVERYLKQOQLLGFPGCSGKLICTTAVPNWASWSNKT 598  
DB 560 IEAQOHLIQLTQWGIKQIARILAVERYLKQOQLLGFPGCSGKLICTTAVPNWASWSNKS 619  
QY 599 LDOIWNMTWMEWDEIDNTHLYTLIEESQOQKNOQELLDKQWASL 649  
DB 620 LDKIWNMTWMEWDEIDNTHLYTLIEESQOQKNOQELLDKQWASL 670

RESULT 15  
US-09-134-075-30  
; Sequence 30, Application US/09134075  
; Patent No. 6042836  
; GENERAL INFORMATION:  
; APPLICANT: Berman, Phillip W.  
; APPLICANT: Nakamura, Gerald R.  
; TITLE OF INVENTION: HIV Envelope Polypeptides  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McCutchen, Doyle, Brown & Enersen, LLP  
; STREET: 3 Embarcadero Center  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fast-SEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/134,075  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/448,603  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halliday, Emily  
; REGISTRATION NUMBER: 38903  
; REFERENCE/DOCKET NUMBER: 14918-704  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-393-2000  
; TELEFAX: 415-393-2286  
; TELEX:



; INFORMATION FOR SEQ ID NO: 30:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 857 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-09-134-075-30

```
Query Match      80.9%; Score 2809.5; DB 3; Length 857;
Best Local Similarity 81.3%; Pred. No. 7.9e-224;
Matches 529; Conservative 45; Mismatches 60; Indels 17; Gaps 6;

QY      3 LWTVIYGVGVWKEATITLFCASDAKAYDTEAHNWATHACVPTNPPOEVVLENTENF 62
Db      33 LWTVIYGVGVWKEATITLFCASDAKAYDTEAHNWATHACVPTNPPOEVVLENTENF 92
QY      63 NWNKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTSLIIVWVEQ 122
Db      93 NWNKNNVVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTD--AGNTTNTNSS----REK 146
QY      123 RKGEMRNCSPNITISIRDKVOREVALFYKLDVEPIDDK--KNTTNNTKYRLINCNTSVI 180
Db      147 LEKGIRKNCSENIITTSVRDKMKETALEFNKLDIVPIDDDDRNSTRNSTNYRLISCNTSVI 206
QY      181 TCACPKVSEFPIPIHYCTPTGFALLKCNCKXKNGTGPCTNVSTVCTGTHGIRPVVSTQLLL 240
Db      207 TCACPKVSEFPIPIHYCTPTGFALLKCNCKXKNGTGPCTNVSTVCTGTHGIRPVVSTQLLL 266
QY      241 NGSLEAEVWIRSENFNTNAKTIIVQLNVSVSEINCTRPNNHTRKRVTLGPGRVWYTTGEI 300
Db      267 NGSLEAEVWIRSENFNTNAKTIIVQLTEPVKINCTRPNNHTRKRSIPIGPGRAFVATGDI 326
QY      301 LGNTRQAHNCSRAQWNTLQIATILREDFGNKTIAFNQSSGGDPEIVMHSFNCGGEPFF 360
Db      327 LGNTRQAHNCSRAQWNTLQIATILREDFGNKTIIFNHSSGGDPEIVMHSFNCRGEFF 386
QY      361 YCNSTQLFNSAWNT--SNGTWSVTRKOKDGTGDIITLPCRIRKQIINRCVGVGKAMYALPI 418
Db      387 YCNSTQLFNSAWNT--SNGTWSVTRKOKDGTGDIITLPCRIRKQIINRCVGVGKAMYALPI 440
QY      419 KGLIRCSSNITGLLTRDGGNGENQTTETPRPGGDMRDNWRSELYKYVVKIEPLGVAPT 478
Db      441 RQGIKRCSSNITGLLTRDGGNSNMNETPRPGGDMRDNWRSELYKYVVKIEPLGVAPT 500
QY      479 KAKRVVQREKRAVGMLGAMFLGFLGAAGSTVGAATSMALTVOARQLLSGIVQQNNLLRA 538
Db      501 KAKRVVQREKRAVG--IGAVFLGFLGAAGSTVGAASITLTVOARQLLSGIVQQNNLLRA 559
QY      539 IKAQQHLLQLTWGIGKQLQARILAVERYLKQQLLGFWGCGSKLICCTTAVPWNASWSNKT 598
Db      560 IKAQQHLLQLTWGIGKQLQARILAVERYLKQQLLGFWGCGSKLICCTTAVPWNASWSNKS 619
QY      599 LDQIWNNTMWEWDREIDNYTHLIYTLIEESQNOQEKQOELLQDLQWASL 649
Db      620 LDQIWNNTMWEWDREIEIENYTHLIYTLIEESQNOQEKQOELLQDLQWASL 670
```

Search completed: January 13, 2004, 09:12:33

Job time : 24 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2004, 09:10:37 ; Search time 38 Seconds  
(without alignments)  
3441.574 Million cell updates/sec

Title: US-09-938-406-1\_COPY\_33\_681

Perfect score: 3474

Sequence: 1 ANLWTVYGVVPMKEATT.....QNQKKNQQLQLDKWASL 649

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	3474	100.0	868	10	US-09-938-406-1
2	2883.5	83.0	842	12	US-10-190-435-2
3	2883.5	83.0	842	12	US-10-241-009-2
4	2883.5	83.0	842	12	US-10-190-434B-2
5	2883.5	83.0	842	12	US-10-190-305A-2
6	2883.5	83.0	847	10	US-09-476-242-2
7	2827.5	81.4	643	15	US-10-032-162-13
8	2809.5	80.9	856	10	US-09-476-242-1
9	2795.5	80.5	856	15	US-10-196-515-11
10	2795.5	80.5	861	15	US-10-026-741-103
11	2757	79.4	726	15	US-10-196-515-3
12	2755	79.3	759	15	US-10-196-515-12
13	2728.5	78.5	619	10	US-09-891-609-4
14	2728.5	78.5	646	10	US-09-891-609-2
15	2720.5	78.3	625	15	US-10-032-162-17

16	2585	74.4	860	12	US-10-190-435-6	Sequence 6, Appli
17	2585	74.4	860	12	US-10-241-009-6	Sequence 6, Appli
18	2585	74.4	860	12	US-10-190-434B-6	Sequence 6, Appli
19	2585	74.4	860	12	US-10-190-305A-6	Sequence 6, Appli
20	2577.5	74.2	853	12	US-10-369-294-13	Sequence 13, Appl
21	2566	73.9	855	12	US-10-190-435-144	Sequence 144, App
22	2563.5	73.8	858	12	US-10-190-435-150	Sequence 150, App
23	2557.5	73.6	579	15	US-10-032-162-15	Sequence 15, Appl
24	2553	73.5	870	12	US-10-190-435-127	Sequence 127, App
25	2552.5	73.5	867	12	US-10-190-435-3	Sequence 3, Appli
26	2552.5	73.5	867	12	US-10-190-435-126	Sequence 126, App
27	2552.5	73.5	867	12	US-10-241-009-3	Sequence 3, Appli
28	2552.5	73.5	867	12	US-10-190-434B-3	Sequence 3, Appli
29	2552.5	73.5	867	12	US-10-190-305A-3	Sequence 3, Appli
30	2551	73.4	855	12	US-10-369-294-11	Sequence 11, Appl
31	2551	73.4	855	12	US-10-369-294-12	Sequence 12, Appl
32	2550	73.4	845	12	US-10-190-435-129	Sequence 129, App
33	2550	73.4	845	12	US-10-190-435-130	Sequence 130, App
34	2548.5	73.4	869	12	US-10-190-435-4	Sequence 4, Appli
35	2548.5	73.4	869	12	US-10-241-009-4	Sequence 4, Appli
36	2548.5	73.4	869	12	US-10-190-434B-4	Sequence 4, Appli
37	2548.5	73.4	869	12	US-10-190-305A-4	Sequence 4, Appli
38	2542.5	73.2	861	12	US-10-190-435-139	Sequence 139, App
39	2541	73.1	865	12	US-10-190-435-140	Sequence 140, App
40	2531	72.9	857	12	US-10-190-435-138	Sequence 138, App
41	2529	72.8	862	12	US-10-190-435-142	Sequence 142, App
42	2528.5	72.8	844	10	US-09-991-258-19	Sequence 19, Appl
43	2524.5	72.7	862	12	US-10-190-435-141	Sequence 141, App
44	2523	72.6	845	12	US-10-190-435-143	Sequence 143, App
45	2519	72.5	803	12	US-10-190-435-134	Sequence 134, App

ALIGNMENTS

RESULT 1

US-09-938-406-1  
; Sequence 1, Application US/09938406  
; Patent No. US20020155120A1  
; GENERAL INFORMATION:  
; APPLICANT: Vancott, George  
; APPLICANT: Birx, Deborah  
; APPLICANT: Vancott, Thomas  
; TITLE OF INVENTION: PROTEIN AND PEPTIDE VACCINES FOR  
; TITLE OF INVENTION: INDUCING MUCOSAL IMMUNITY  
; FILE REFERENCE: 40646-20002.10  
; CURRENT APPLICATION NUMBER: US/09/938,406  
; CURRENT FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 09/214,701  
; PRIOR FILING DATE: 1999-09-30  
; PRIOR APPLICATION NUMBER: PCT/US 97/12253  
; PRIOR FILING DATE: 1997-07-10  
; PRIOR APPLICATION NUMBER: US 60/021,687  
; PRIOR FILING DATE: 1996-07-10  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 868  
; TYPE: PRT  
; ORGANISM: Virus HIV-1  
US-09-938-406-1

Query Match	100.0%;	Score 3474;	DB 10;	Length 868;
Best Local Similarity	100.0%;	Pred. No. 5.4e-306;		
Matches 649;	Conservative	0;	Mismatches	0;
	Indels	0;	Gaps	0;
QY	1	ANLWTVYGVVPMKEATTTLFCASDAKAYTEAHNWTACHCPTNPDPQEVLENVTE	60	
DB	33	ANLWTVYGVVPMKEATTTLFCASDAKAYTEAHNWTACHCPTNPDPQEVLENVTE	92	
QY	61	NFNWKNWVQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTNTTSLIIVW	120	
DB	93	NFNWKNWVQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTNTTSLIIVW	152	

121 EORKGEMRNCNFNITTSIRDKVOREYALFYKLDVEPIDDNKNTNTNTKYRLINCNTSVI 180  
153 EORKGEMRNCNFNITTSIRDKVOREYALFYKLDVEPIDDNKNTNTNTKYRLINCNTSVI 212  
181 TOACPKVSFEPIPIHYCTPTGPFALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTOLL 240  
213 TOACPKVSFEPIPIHYCTPTGPFALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTOLL 272  
241 NGSLAEVEVIRSENFNNAKTIIIVQLNVSEINCTRENHTRKRVTLGPGRVWYTTGEI 300  
273 NGSLAEVEVIRSENFNNAKTIIIVQLNVSEINCTRENHTRKRVTLGPGRVWYTTGEI 332  
301 LGNIROAHCHNISRAQWNTLQOIAATTLREQFGNKTIAFNQSSGGDPEIWMHSFNCGGEFF 360  
333 LGNIROAHCHNISRAQWNTLQOIAATTLREQFGNKTIAFNQSSGGDPEIWMHSFNCGGEFF 392  
361 YCNSTOLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRKQIINRWQVVGKAMVALPIKG 420  
393 YCNSTOLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRKQIINRWQVVGKAMVALPIKG 452  
421 LIRCSSNITGLLLTRDGGENQTTIEFRPGGDMRDNRSELYKYKVVKIEPLGVAPTKA 480  
453 LIRCSSNITGLLLTRDGGENQTTIEFRPGGDMRDNRSELYKYKVVKIEPLGVAPTKA 512  
481 KRVVQREKRAVGMGLGAMFLGFLGAAGTSMALTVOARQLLSGIVQOONLLRAIK 540  
513 KRVVQREKRAVGMGLGAMFLGFLGAAGTSMALTVOARQLLSGIVQOONLLRAIK 572  
541 AQOHLQLTVMGKIQLOARILAVERYLKDQQLLGFPGGSGKLICTTAVPWNASWNTLID 600  
573 AQOHLQLTVMGKIQLOARILAVERYLKDQQLLGFPGGSGKLICTTAVPWNASWNTLID 632  
601 QIWNMTWMDREIDNYTHLYTLIESQOQEKNOQELLQLDKWSL 649  
633 QIWNMTWMDREIDNYTHLYTLIESQOQEKNOQELLQLDKWSL 681

## RESULT 2

US-10-190-435-2

; Sequence 2, Application US/10190435

; Publication No. US20030143248A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: BARNETT, Susan W.

; APPLICANT: LIAN, Ying

; APPLICANT: ENGELBRECHT, Susan

; APPLICANT: VAN RENSBURG, Estrelita J.

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C

; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

; FILE REFERENCE: PP18133.003 / 2302-18133

; CURRENT APPLICATION NUMBER: US/10/190,435

; CURRENT FILING DATE: 2002-12-30

; NUMBER OF SEQ ID NOS: 319

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 842

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SFI62

US-10-190-435-2

## Query Match

Best Local Similarity 83.0%; Score 2883.5; DB 12; Length 842;

Matches 544; Conservative 40; Mismatches 38; Indels 31; Gaps 7;

QY

3 LWVTYVYGVPMWKEATTTIFCASDAKAYDTEAHNWNWATHACVPTNPQEVLENVTENF 62

DB

28 LWVTYVYGVPMWKEATTTIFCASDAKAYDTEAHNWNWATHACVPTNPQEVLENVTENF 87

QY

63 NWKKNVVEQMHEDIISLDQSLKPCVKLTPLCVLTNCTDLNNTNTTETLSIIWVWQ 122

DB

88 NWKKNVVEQMHEDIISLDQSLKPCVKLTPLCVLTNCTDLNNTNTKSSN-----WKE 141

88 NWKKNVVEQMHEDIISLDQSLKPCVKLTPLCVLTNCTDLNNTNTKSSN-----WKE 141  
123 RGKEMRNCNFNITTSIRDKVOREYALFYKLDVEPIDDNKNTNTNTKYRLINCNTSVI 182  
142 MDREIKNCSPFKVTTISRNRKQREYALFYKLDVVVIDN-----DNTSYKLINCNTSVITQ 196  
183 ACPKVSFEPIPIHYCTPTGPFALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTOLL 242  
197 ACPKVSFEPIPIHYCAPAGFAILKNDKFKNGSGPCTNVSTVQCTHGIRPVVSTOLL 256  
243 SLAEVEVIRSENFNNAKTIIIVQLNVSEINCTRENHTRKRVTLGPGRVWYTTGEILG 302  
257 SLAEVEVIRSENFNNAKTIIIVQLKESVEINCTRENHTRKRVTLGPGRVWYTTGEILG 316  
303 NIROAHCHNISRAQWNTLQOIAATTLREQFGNKTIAFNQSSGGDPEIWMHSFNCGGEFF 362  
317 DIROAHCHNISRAQWNTLQOIAATTLREQFGNKTIAFNQSSGGDPEIWMHSFNCGGEFF 376  
363 NSTOLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRKQIINRWQVVGKAMVALP 417  
377 NSTOLFNSAWNTSNGTWSVTRKQDGTGDIITLPCRKQIINRWQVVGKAMVALP 423  
418 IKGLIRCSSNITGLLLTRDGGENQTTIEFRPGGDMRDNRSELYKYKVVKIEPLGV 476  
424 IRGQIRCSSNITGLLLTRDGGENQTTIEFRPGGDMRDNRSELYKYKVVKIEPLGV 483  
477 PTKAKRRVQREKRAVGMGLGAMFLGFLGAAGTSMALTVOARQLLSGIVQOONLL 536  
484 PTKAKRRVQREKRAVGMGLGAMFLGFLGAAGTSMALTVOARQLLSGIVQOONLL 542  
537 RAIKAQOHLQLTVMGKIQLOARILAVERYLKDQQLLGFPGGSGKLICTTAVPWNASW 596  
543 RAIKAQOHLQLTVMGKIQLOARILAVERYLKDQQLLGFPGGSGKLICTTAVPWNASW 602  
597 KTLDOIWNNMTWMDREIDNYTHLYTLIESQOQEKNOQELLQLDKWSL 649  
603 KSLDOIWNNMTWMDREIDNYTHLYTLIESQOQEKNOQELLQLDKWSL 655

## RESULT 3

US-10-241-009-2

; Sequence 2, Application US/10241009

; Publication No. US20030170614A1

; GENERAL INFORMATION:

; APPLICANT: ZUR MEGEDE, Jan

; APPLICANT: BARNETT, Susan

; APPLICANT: LIAN, Ying

; APPLICANT: VAN RENSBURG, Estrelita J.

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B

; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

; FILE REFERENCE: 2300-1621.21

; CURRENT APPLICATION NUMBER: US/10/241,009

; CURRENT FILING DATE: 2002-12-13

; NUMBER OF SEQ ID NOS: 68

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 842

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SFI62

US-10-241-009-2

## Query Match

Best Local Similarity 83.0%; Score 2883.5; DB 12; Length 842;

Matches 544; Conservative 40; Mismatches 38; Indels 31; Gaps 7;

QY

3 LWVTYVYGVPMWKEATTTIFCASDAKAYDTEAHNWNWATHACVPTNPQEVLENVTENF 62

DB

28 LWVTYVYGVPMWKEATTTIFCASDAKAYDTEAHNWNWATHACVPTNPQEVLENVTENF 87

QY

63 NWKKNVVEQMHEDIISLDQSLKPCVKLTPLCVLTNCTDLNNTNTTETLSIIWVWQ 122

DB

88 NWKKNVVEQMHEDIISLDQSLKPCVKLTPLCVLTNCTDLNNTNTKSSN-----WKE 141

QY 123 RKGEMRNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYBLINCNTSVITQ 182  
Db 142 MRGEIKNCSEFKVTTTSIRKMKQKEYALFYKLDVVPIDN-----DNTSYKLINCNTSVITQ 196  
QY 183 ACPKVSFEPIPHYCTPTGTFALLKCKNDKXFNCTVSTVQCTHGIKIRPVVSTQLLNG 242  
Db 197 ACPKVSFEPIPHYCAPAGFAILKCKNDKXFNCTVSTVQCTHGIKIRPVVSTQLLNG 256  
QY 243 SLAEVWIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPGKRVWYTTGEILG 302  
Db 257 SLAEGWIRSENFNTNAKTIIVQLKESVEINCTRPNNHTRKRSITIGPGRAFVATGDIIG 316  
QY 303 NTRQAHNCSRAQWNTLQOIATTLREQGNKTIAPNQSSGGDPETVWMSFNCGGEFFYC 362  
Db 317 DIRQAHNCSISGKWNNTLQOIVTKLQAGFNKTIIVFKQSSGGDPETVWMSFNCGGEFFYC 376  
QY 363 NSTQLFNSAWNT-----SNGTWSVTRKQKDTGDIITLPCRKQIINRWQVVGKAMYALP 417  
Db 377 NSTQLFNSWNTNNTIGPNTNGT-----ITLPCRKQIINRWQVVGKAMYALP 423  
QY 418 IKGLRCSNITGLLLTRDGGGE-NOTTEIFRPGGDMRDNRSELYKYKVKIEPLGVA 476  
Db 424 IRGQIRCSNITGLLLTRDGGGEISNTTIFRPGGDMRDNRSELYKYKVKIEPLGVA 483  
QY 477 PTKAKRRVVQREKRAVGMGLGAMFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOQNILL 536  
Db 484 PTKAKRRVVQREKRAV-TLGAMFLGFLGAAGSTMGARSULTVQARQLLSGIVQOQNILL 542  
QY 537 RAIKAQHLQLTQVWGIKQARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASWN 596  
Db 543 RAIKAQHLQLTQVWGIKQARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASWN 602  
QY 597 KTLDOQWNNMTWMDREIDNTHLYTLIESQKQEKNOQELLDKXWASL 649  
Db 603 KSLDQWNNMTWMDREIDNTHLYTLIESQKQEKNOQELLDKXWASL 655

## RESULT 4

US-10-190-434B-2  
; Sequence 2, Application US/10190434B  
; Publication No. US20030194800A1

## GENERAL INFORMATION:

; APPLICANT: ZUR MEDEDE, Jan

; APPLICANT: BARNETT, Susan

; APPLICANT: LIAN, Ying

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B

; FILE REFERENCE: 2300-1621-20

; CURRENT APPLICATION NUMBER: US/10/190,434B

; CURRENT FILING DATE: 2002-07-05

; NUMBER OF SEQ ID NOS: 68

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 842

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SF162

US-10-190-434B-2

## Query Match

Best Local Similarity 83.0%; Score 2883.5; DB 12; Length 842;

Matches 544; Conservative 40; Mismatches 38; Indels 31; Gaps 7;

QY 3 LWTVYVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 62  
Db 28 LWTVYVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 87  
QY 63 NNMKNMVEQMHEDIISLWDSKPCVKLTPLCVTLNCTDLNNTNTTETLSIIVWQ 122  
Db 88 NNMKNMVEQMHEDIISLWDSKPCVKLTPLCVTLNCTDLNNTNTKSSN----WKE 141

QY 123 RKGEMRNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYBLINCNTSVITQ 182  
Db 142 MRGEIKNCSEFKVTTTSIRKMKQKEYALFYKLDVVPIDN-----DNTSYKLINCNTSVITQ 196  
QY 183 ACPKVSFEPIPHYCTPTGTFALLKCKNDKXFNCTVSTVQCTHGIKIRPVVSTQLLNG 242  
Db 197 ACPKVSFEPIPHYCAPAGFAILKCKNDKXFNCTVSTVQCTHGIKIRPVVSTQLLNG 256  
QY 243 SLAEVWIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPGKRVWYTTGEILG 302  
Db 257 SLAEGWIRSENFNTNAKTIIVQLKESVEINCTRPNNHTRKRSITIGPGRAFVATGDIIG 316  
QY 303 NTRQAHNCSRAQWNTLQOIATTLREQGNKTIAPNQSSGGDPETVWMSFNCGGEFFYC 362  
Db 317 DIRQAHNCSISGKWNNTLQOIVTKLQAGFNKTIIVFKQSSGGDPETVWMSFNCGGEFFYC 376  
QY 363 NSTQLFNSAWNT-----SNGTWSVTRKQKDTGDIITLPCRKQIINRWQVVGKAMYALP 417  
Db 377 NSTQLFNSWNTNNTIGPNTNGT-----ITLPCRKQIINRWQVVGKAMYALP 423  
QY 418 IKGLRCSNITGLLLTRDGGGE-NOTTEIFRPGGDMRDNRSELYKYKVKIEPLGVA 476  
Db 424 IRGQIRCSNITGLLLTRDGGGEISNTTIFRPGGDMRDNRSELYKYKVKIEPLGVA 483  
QY 477 PTKAKRRVVQREKRAVGMGLGAMFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOQNILL 536  
Db 484 PTKAKRRVVQREKRAV-TLGAMFLGFLGAAGSTMGARSULTVQARQLLSGIVQOQNILL 542  
QY 537 RAIKAQHLQLTQVWGIKQARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASWN 596  
Db 543 RAIKAQHLQLTQVWGIKQARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASWN 602  
QY 597 KTLDOQWNNMTWMDREIDNTHLYTLIESQKQEKNOQELLDKXWASL 649  
Db 603 KSLDQWNNMTWMDREIDNTHLYTLIESQKQEKNOQELLDKXWASL 655

## RESULT 5

US-10-190-305A-2

; Sequence 2, Application US/10190305A

; Publication No. US20030198621A1

## GENERAL INFORMATION:

; APPLICANT: ZUR MEDEDE, Jan

; APPLICANT: BARNETT, Susan

; APPLICANT: LIAN, Ying

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE B AND/OR

; FILE REFERENCE: 2302-18702 / 18702.002

; CURRENT APPLICATION NUMBER: US/10/190,305A

; CURRENT FILING DATE: 2002-07-05

; NUMBER OF SEQ ID NOS: 93

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 842

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: SF162

US-10-190-305A-2

## Query Match

Best Local Similarity 83.0%; Score 2883.5; DB 12; Length 842;

Matches 544; Conservative 40; Mismatches 38; Indels 31; Gaps 7;

QY 3 LWTVYVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 62  
Db 28 LWTVYVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQEVLENVTENF 87  
QY 63 NNMKNMVEQMHEDIISLWDSKPCVKLTPLCVTLNCTDLNNTNTTETLSIIVWQ 122  
Db 88 NNMKNMVEQMHEDIISLWDSKPCVKLTPLCVTLNCTDLNNTNTKSSN----WKE 141  
QY 123 RKGEMRNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYBLINCNTSVITQ 182

```
Db 142 MDRGEIKNCSFKVTTIRNMQKEALFYKLDVWPIDN-----DNTSVKLNINCNSVITQ 196
Qy 183 ACPEVSEPIPIHYCTPTGFAILKCNCKKFGNGTGPCTNVSTVQCTHGIRPVVSTQILLNG 242
Db 197 ACPEVSEPIPIHYCAPAGFAILKCNCKKFGNGTGPCTNVSTVQCTHGIRPVVSTQILLNG 256
Qy 243 SLABEEVVRSENFTNNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVWYITGEILG 302
Db 257 SLABEGVVRSENFTNNAKTIIVQLKESVEINCTRPNNHTRKSTITIGGRAFYATGDIIG 316
Qy 303 NIROAHCNISRAQNNNTLQOATTLREOFNGKTIARNOSSGGDPEIVMHSFNCGGERFYC 362
Db 317 DIROAHCNISGEKWNNTLKQIVTKLQAGFGNKTIVFKOSSGGDPEIVMHSFNCGGERFYC 376
Qy 363 NSTQLFNSAMNVT-----SNGTWSVTRKQKOTGDIIITLPCRKIQIINRWQVVGKAMYALP 417
Db 377 NSTQLFNSAMNVTGPNNTNGT-----ITLPCRKIQIINRWQVVGKAMYAPP 423
Qy 418 IKGLIRCSSNTIGLLTRDGGGE-NQTEIPRPGGDMRDNRSELKYKVKVIEPLGVA 476
Db 424 IRGQIRCSSNTIGLLTRDGGGEISNTTEIPRPGGDMRDNRSELKYKVKVIEPLGVA 483
Qy 477 PTKAKRRVVQREKRAVGMGLGAMFELGAGSTWGATSMALTVOARQLLSGIVQOQNLL 536
Db 484 PTKAKRRVVQREKRAV-TLGAFLGFLGAGSTWGASLTTLTVQARQLLSGIVQOQNLL 542
Qy 537 RAIKAQOHLQLTVWGIKQLOARILAVERYLKDOQLLGFWSGKLICTTAVPWNASWSN 596
Db 543 RAIKAQOHLQLTVWGIKQLOARILAVERYLKDOQLLGFWSGKLICTTAVPWNASWSN 602
Qy 597 KTLQIWNMTWMEWDREIDNYTHLIYTLIBESQOQEKNOQELLQDKWASL 649
Db 603 KSLDQIWNMTWMEWDREIDNYTHLIYTLIBESQOQEKNOQELLQDKWASL 655

RESULT 6
US-09-476-242-2
; Sequence 2, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Karin
; APPLICANT: MARTIN, Eric
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476,242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-476-242-2

Query Match 83.0%; Score 2883.5; DB 10; Length 847;
Best Local Similarity 83.3%; Pred. No. 2e-252;
Matches 544; Conservative 40; Mismatches 38; Indels 31; Gaps 7;

Qy 3 LWVTYYGVVPWKATITLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 62
Db 33 LWVTYYGVVPWKATITLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 92
Qy 63 NWNKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELIIIVWMEQ 122
Db 93 NWNKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTKSSN-----WKE 146
Qy 123 RKGEMRNCSEFNITTSIRDKQREYALFYKLDVPEIDNKNNTNNTKYLRLNCNTSVITQ 182
Db 147 MDRGEIKNCSFKVTTIRNMQKEALFYKLDVWPIDN-----DNTSVKLNINCNSVITQ 201
Qy 183 ACPEVSEPIPIHYCTPTGFAILKCNCKKFGNGTGPCTNVSTVQCTHGIRPVVSTQILLNG 242
```

```
Db 202 ACPEVSEPIPIHYCAPAGFAILKCNCKKFGNGTGPCTNVSTVQCTHGIRPVVSTQILLNG 261
Qy 243 SLABEEVVRSENFTNNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVWYITGEILG 302
Db 262 SLABEGVVRSENFTNNAKTIIVQLKESVEINCTRPNNHTRKSTITIGGRAFYATGDIIG 321
Qy 303 NIROAHCNISRAQNNNTLQOATTLREOFNGKTIARNOSSGGDPEIVMHSFNCGGERFYC 362
Db 322 DIROAHCNISGEKWNNTLKQIVTKLQAGFGNKTIVFKOSSGGDPEIVMHSFNCGGERFYC 381
Qy 363 NSTQLFNSAMNVT-----SNGTWSVTRKQKOTGDIIITLPCRKIQIINRWQVVGKAMYALP 417
Db 382 NSTQLFNSAMNVTGPNNTNGT-----ITLPCRKIQIINRWQVVGKAMYAPP 428
Qy 418 IKGLIRCSSNTIGLLTRDGGGE-NQTEIPRPGGDMRDNRSELKYKVKVIEPLGVA 476
Db 429 IRGQIRCSSNTIGLLTRDGGGEISNTTEIPRPGGDMRDNRSELKYKVKVIEPLGVA 488
Qy 477 PTKAKRRVVQREKRAVGMGLGAMFELGAGSTWGATSMALTVOARQLLSGIVQOQNLL 536
Db 489 PTKAKRRVVQREKRAV-TLGAFLGFLGAGSTWGASLTTLTVQARQLLSGIVQOQNLL 547
Qy 537 RAIKAQOHLQLTVWGIKQLOARILAVERYLKDOQLLGFWSGKLICTTAVPWNASWSN 596
Db 548 RAIKAQOHLQLTVWGIKQLOARILAVERYLKDOQLLGFWSGKLICTTAVPWNASWSN 607
Qy 597 KTLQIWNMTWMEWDREIDNYTHLIYTLIBESQOQEKNOQELLQDKWASL 649
Db 608 KSLDQIWNMTWMEWDREIDNYTHLIYTLIBESQOQEKNOQELLQDKWASL 660

RESULT 7
US-10-032-162-13
; Sequence 13, Application US/10032162
; Publication No. US20030052839A1
; GENERAL INFORMATION:
; APPLICANT: BINLEY, JAMES M
; APPLICANT: SCHUELE, NORBERT
; APPLICANT: OLSON, WILLIAM C
; APPLICANT: PAUL, MADDOX J
; APPLICANT: JOHN, MOORE P
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF
; FILE REFERENCE: 2048/59331az
; CURRENT APPLICATION NUMBER: US/10/032,162
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-032-162-13

Query Match 81.4%; Score 2827.5; DB 15; Length 643;
Best Local Similarity 82.2%; Pred. No. 1.6e-247;
Matches 532; Conservative 40; Mismatches 56; Indels 19; Gaps 4;

Qy 3 LWVTYYGVVPWKATITLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 62
Db 4 LWVTYYGVVPWKATITLFCASDAKAYDTEAHNVWATHACVPTNPQEVLENVTENF 63
Qy 63 NWNKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELIIIVWMEQ 122
Db 64 NWNKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTDSEGT----- 117
Qy 123 RKGEMRNCSEFNITTSIRDKQREYALFYKLDVPEIDNKNNTNNTKYLRLNCNTSVITQ 182
Db 118 -ERGEMRNCSEFNITTSIRDEVOKEALFYKLDVPEIDN-----NNTSVRLISCDTSVITQ 171
Qy 183 ACPEVSEPIPIHYCTPTGFAILKCNCKKFGNGTGPCTNVSTVQCTHGIRPVVSTQILLNG 242
```

```
Db 172 ACPISEPIPIHYCAPAGFAILACNDKTFNGKPCVKSVQCTHGRPVVSTQLLNG 231
Qy 243 SLAEEVWIRSENFTNAKTIIVOLNVSVEINCTRPNNHTKRVTLGPRVWYTTGELG 302
Db 232 SLAEEVWIRSDNFTNAKTIIVOLKBSVEINCTRPNNHTKRSIHIGPRAFYTTGELG 291
Qy 303 NROAHNCISRAQNNLTQOATTLREQFGNKTIAFNOSGGDPEIVMHSNCGGEFFYC 362
Db 292 DIRAHNCISRAKNDTLKQIVLKRQFENKTIVFNHSSGGDPEIVMHSNCGGEFFYC 351
Qy 363 NSTOLFNSAMNVTSGTWSVTRKOKDTGDIITLPCRIKQIINRWQVVGKAMAYALPIKGLI 422
Db 352 NSTOLFNSAMNVTSGTWSVTRKOKDTGDIITLPCRIKQIINRWQVVGKAMAYALPIKGLI 405
Qy 423 RCSSNITGLLTRDGGGEGNQTTEIFRPGGDMRDNRSSELYKYKVKIEPLGVAPTAKR 482
Db 406 RCSSNITGLLTRDGGGEGNQTTEIFRPGGDMRDNRSSELYKYKVKIEPLGVAPTAKR 465
Qy 483 RVVOREKRAVGMFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOQNNLLRAIKAO 542
Db 466 RVVOREKRAVG-IGAVFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOQNNLLRAIKAO 524
Qy 543 QHLLQTLTVMGKIQARILAVERYLKDQQLLGFPGWCSGKLICTTAVPWNASWSNKTLDQI 602
Db 525 QHLLQTLTVMGKIQARILAVERYLKDQQLLGFPGWCSGKLICTTAVPWNASWSNKTLDQI 584
Qy 603 WNNMTWMEWDREIDNYTHLYTLIEESQOQEKNOQELLQDKWASL 649
Db 585 WNNMTWMEWDREIDNYTHLYTLIEESQOQEKNOQELLQDKWASL 631

RESULT 8
US-09-476-242-1
; Sequence 1, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Karin
; APPLICANT: MARTIN, Eric
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476,242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-476-242-1

Query Match 80.9%; Score 2809.5; DB 10; Length 856;
Best Local Similarity 82.6%; Pred. No. 1.1e-245;
Matches 538; Conservative 38; Mismatches 56; Indels 19; Gaps 9;
Qy 3 LWTVYGVGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPNPQVVLNVTENF 62
Db 34 LWTVYGVGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPNPQVVLNVTENF 93
Qy 63 NMKNKMWQEHEDIISLWDSQSLKPCVKLTPLCVTLNCTDLNNTNTNITELSIIVVWEQ 122
Db 94 NMKNKMWQEHEDIISLWDSQSLKPCVKLTPLCVTLNCTDLNNTNTSSSGRIME- 150
Qy 123 RGKGMNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNNTNTNTKYRLINCNITSVITQ 182
Db 151 --KGEIKNCSFNISTSRGKVQKEYAFFYKLDIIPDN-----DTSYKLTSCNTSVITQ 203
Qy 183 ACPKVSFEPIPIHYCTPTGFAILLKNDKFNCTGCTNVTSTVQCTHGIRPVVSTQLLNG 242
Db 204 ACPKVSFEPIPIHYCAPAGFAILKCNKNTFNGTGPCTNVTSTVQCTHGIRPVVSTQLLNG 263
Qy 243 SLAEEVWIRSENFTNAKTIIVOLNVSVEINCTRPNNHTKRVTL--GPGRVWYTTGEL 300
```

```
Db 264 SLAEEVWIRSVNFTDNAKTIIVOLNVSVEINCTRPNNHTKRIORGPGRFVTIGKI 323
Qy 301 LGNTRQAHNCISRAQNNLTQOATTLREQFG-NKTIAFNOSGGDPEIVMHSNCGGEF 359
Db 324 -GNMROAHNCISRAKNNLTQIASKLREQFGNNTKTIIFKSSGGDPEIVTHSNCGGEF 382
Qy 360 FYCNSTOLFNSAMNVTSGTWSVTRKOKDTG-DIITLPCRIKQIINRWQVVGKAMAYALPI 418
Db 383 FYCNSTOLFNSAMNVTSGTWSVTRKOKDTG-DIITLPCRIKQIINRWQVVGKAMAYALPI 439
Qy 419 KGLRCSSNITGLLTRDGGGEGNQTTEIFRPGGDMRDNRSSELYKYKVKIEPLGVAPT 478
Db 440 SGQIRCSNITGLLTRDGGGEGNQTTEIFRPGGDMRDNRSSELYKYKVKIEPLGVAPT 499
Qy 479 KAKERVVQREKRAVGMFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOQNNLLRA 538
Db 500 KAKERVVQREKRAVG-IGALFLGFLGAAGSTMGTSMALTVOARQLLSGIVQOQNNLLRA 558
Qy 539 IKAQHLLQTLTVMGKIQARILAVERYLKDQQLLGFPGWCSGKLICTTAVPWNASWSNKT 598
Db 559 IEAQHLLQTLTVMGKIQARILAVERYLKDQQLLGFPGWCSGKLICTTAVPWNASWSNKS 618
Qy 599 LDQIWNMTWMEWDREIDNYTHLYTLIEESQOQEKNOQELLQDKWASL 649
Db 619 LEQIWNHTTWMEWDREIDNYTHLYTLIEESQOQEKNOQELLQDKWASL 669

RESULT 9
US-10-196-515-11
; Sequence 11, Application US/10196515
; Publication No. US20030091594A1
; GENERAL INFORMATION:
; APPLICANT: HOXIE, James A.
; APPLICANT: LABRANCHE, Celia C.
; APPLICANT: DOMS, Robert W.
; APPLICANT: HOFFMAN, Trevor L.
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
; FILE REFERENCE: Hoxie 9596-104U1 (0282)
; CURRENT APPLICATION NUMBER: US/10/196,515
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: US/09/337,387
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/317,556
; PRIOR FILING DATE: 1999-05-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-10-196-515-11

Query Match 80.5%; Score 2795.5; DB 15; Length 856;
Best Local Similarity 82.3%; Pred. No. 2e-244;
Matches 536; Conservative 38; Mismatches 58; Indels 19; Gaps 9;
Qy 3 LWTVYGVGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPNPQVVLNVTENF 62
Db 34 LWTVYGVGVVWKEATTLFCASDAKAYDTEAHNVWATHACVPTNPNPQVVLNVTENF 93
Qy 63 NMKNKMWQEHEDIISLWDSQSLKPCVKLTPLCVTLNCTDLNNTNTNITELSIIVVWEQ 122
Db 94 NMKNKMWQEHEDIISLWDSQSLKPCVKLTPLCVSLKCTDL--KNDTNTSSSGRIME- 150
Qy 123 RGKGMNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNNTNTNTKYRLINCNITSVITQ 182
Db 151 --KGEIKNCSFNISTSRGKVQKEYAFFYKLDIIPDN-----DTSYKLTSCNTSVITQ 203
Qy 183 ACPKVSFEPIPIHYCTPTGFAILLKNDKFNCTGCTNVTSTVQCTHGIRPVVSTQLLNG 242
Db 204 ACPKVSFEPIPIHYCAPAGFAILKCNKNTFNGTGPCTNVTSTVQCTHGIRPVVSTQLLNG 263
```

243 SLABEEVIRSENFNNNAKTIIVQLNVSVINCTPPNNHTRKSVTL--GPGRVWYTTGEI 300  
264 SLABEEVIRSENFNNNAKTIIVQLNVSVINCTPPNNHTRKSVTL--GPGRVWYTTGEI 323  
301 LGNIRCAHNTSRAQWNNTLOQIATTLREOFG-NKTIAPNQSSGGDPEIVMHSFNGCGEF 359  
324 -GNMQAHCNISRAQWNNTLOQIATTLREOFG-NKTIAPNQSSGGDPEIVMHSFNGCGEF 382  
360 FYCNSSTQLFNSAWNTSNGTWSVTRKQDGTG-DIITLPCRIKQIINRMQVGVKAMYPALPI 418  
383 FYCNSSTQLFNSAWNTSNGTWSVTRKQDGTG-DIITLPCRIKQIINRMQVGVKAMYPALPI 439  
419 KGLIRCSSNITGLLTRDGGNGNQTTEIPRPGGDMRDNRSELYKYKVKIEPIGVAPT 478  
440 SQIRCSSNITGLLTRDGGNGNQTTEIPRPGGDMRDNRSELYKYKVKIEPIGVAPT 499  
479 KAKRRVOREKRAVGMGLGAGSTMGATSMALTQVQARQLLSGIVQOQNNLLRA 538  
500 KAKRRVOREKRAVGMGLGAGSTMGATSMALTQVQARQLLSGIVQOQNNLLRA 558  
539 IKAQOHLQLTWVGIKQLQARILAVERYLKDQQLLGFNGCSGKLICTTAVPWNASWSNKT 598  
559 IEAQOHLQLTWVGIKQLQARILAVERYLKDQQLLGFNGCSGKLICTTAVPWNASWSNKT 618  
599 LQIWNMTWMEWDREIDNYTHLYTLIESQNOQEKNOQELLDKWSL 649  
619 LQIWNMTWMEWDREIDNYTHLYTLIESQNOQEKNOQELLDKWSL 669

RESULT 10  
US-10-026-741-103  
Sequence 103, Application US/10026741  
Publication No. US20030049604A1  
GENERAL INFORMATION:  
APPLICANT: CHARNEAU, PIERRE  
BORMAN, ANDREW  
QUILLIENT, CAROLINE  
GUETARD, DENISE  
MONTAGNIER, LUC  
DONJON DE SAINT-MARTIN, JACQUELINE  
COHEN, JACQUES  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF HIV-1 TYPE (OR SUBTYPE) ANTIGENS  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/026,741  
FILING DATE: 27-Dec-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/817,441  
FILING DATE: 31-AUG-1998  
APPLICATION NUMBER: PCT/FR 95/01391  
FILING DATE: 20-OCT-1995  
APPLICATION NUMBER: FR 9412554  
FILING DATE: 20-OCT-1994  
APPLICATION NUMBER: FR 9502526  
FILING DATE: 03-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.

REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 03260.6005-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 103:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 861 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 103:  
US-10-026-741-103  
Query Match 80.5%; Score 2795; DB 15; Length 861;  
Best Local Similarity 82.4%; Pred. No. 2.2e-244;  
Matches 539; Conservative 37; Mismatches 58; Indels 20; Gaps 10;  
3 LMVTYYGVVPWKBAATTLFCASDAKAYDTEAHNWWATHACVPTNPPOEVVLENTENF 62  
34 LMVTYYGVVPWKBAATTLFCASDAKAYDTEAHNWWATHACVPTNPPOEVVLENTENF 93  
63 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLGNATNTSSNTSSGEMM 153  
94 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLGNATNTSSNTSSGEMM 179  
120 MEQRKGEKMRNCSFNITTSIRKQVREYALFYKLDVEPIDDKNTNTTKYLINCNTSV 205  
154 ME---KGEIKKCSFNISTIRGKQVKEYAFFYKLDIIPIDN-----DITSYLTSCNTSV 239  
180 ITQACPKVSFEPIPIHYCTPTGTFALLKCNKDKFNCTGCTNVSTVQCTHIGRPVYSTQLL 265  
206 ITQACPKVSFEPIPIHYCAPAGFALLKCNKTFNGCTNVSTVQCTHIGRPVYSTQLL 297  
240 LNSLAEEVWIRSENFNNNAKTIIVQLNVSVINCTPPNNHTRK--RVTLPGRVWYTT 325  
266 LNSLAEEVWIRSENFNNNAKTIIVQLNVSVINCTPPNNHTRK--RVTLPGRVWYTT 356  
298 GEILNIRQAHCNISRAQWNNTLOQIATTLREOFG-NKTIAPNQSSGGDPEIVMHSFNGC 384  
326 GK1-GNMQAHCNISRAQWNNTLOQIATTLREOFG-NKTIAPNQSSGGDPEIVMHSFNGC 415  
357 GFEFFCNSTQLFNSAWNTSNGTWSVTRKQDGTG-DIITLPCRIKQIINRMQVGVKAMYA 441  
385 GFEFFCNSTQLFNSAWNTSNGTWSVTRKQDGTG-DIITLPCRIKQIINRMQVGVKAMYA 475  
416 LPIKGLIRCSSNITGLLTRDGGNGNQTTEIPRPGGDMRDNRSELYKYKVKIEPIGV 501  
442 PPISGQIRCSSNITGLLTRDGGNGNQTTEIPRPGGDMRDNRSELYKYKVKIEPIGV 535  
476 APTKAKRRVOREKRAVGMGLGAGSTMGATSMALTQVQARQLLSGIVQOQNNLL 560  
502 APTKAKRRVOREKRAVGMGLGAGSTMGATSMALTQVQARQLLSGIVQOQNNLL 595  
536 LRAIKAQOHLQLTWVGIKQLQARILAVERYLKDQQLLGFNGCSGKLICTTAVPWNASWS 620  
561 LRAIKAQOHLQLTWVGIKQLQARILAVERYLKDQQLLGFNGCSGKLICTTAVPWNASWS 649  
596 NKTLDQIWNMTWMEWDREIDNYTHLYTLIESQNOQEKNOQELLDKWSL 674  
621 NKSLEQIWNMTWMEWDREIDNYTHLYTLIESQNOQEKNOQELLDKWSL 699  
RESULT 11  
US-10-196-515-3  
Sequence 3, Application US/10196515  
Publication No. US20030091594A1  
GENERAL INFORMATION:  
APPLICANT: HOKIE, James A.  
APPLICANT: LABRANCHE, Celia C.  
APPLICANT: DOMS, Robert W.  
APPLICANT: HOFFMAN, Trevor L.



; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND  
; TITLE OF INVENTION: THERAPEUTICS  
; FILE REFERENCE: Hoxie 9596-104U1 (0282)  
; CURRENT APPLICATION NUMBER: US/10/196,515  
; CURRENT FILING DATE: 2002-07-16  
; PRIOR APPLICATION NUMBER: US/09/337,387  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: US 09/317,556  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 726  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-10-196-515-3

Query Match 79.4%; Score 2757; DB 15; Length 726;  
Best Local Similarity 81.3%; Pred. No. 4.8e-241;  
Matches 529; Conservative 43; Mismatches 55; Indels 24; Gaps 9;  
QY 3 LWTVVYGVVWKEATTTILFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62  
DB 34 LWTVVYGVVWKEATTTILFCASDAKAYETEVHNVWATHACVPTDNPQEVVLENTNF 93  
QY 63 NWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVLTNCTDLNNTNTTTELSIIIVWVQ 122  
DB 94 NWKNDMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDL--KNDTNTSSSGRIME- 150  
QY 123 RKGEMRNCSEFNTSIRDKVOREYALFKLDVEPDDNKNTNTNNTKYRLNCTSVITQ 182  
DB 151 --KGEIKNCSEFNTSIRDKVOREYALFKLDVEPDDNKNTNTNNTKYRLNCTSVITQ 203  
QY 183 ACPKVSFEPIHYCTPTGTGTFALLKNDKXFNCTGPTCTNVSTVQCTHGIRPVVSTQLLNG 242  
DB 204 ACPKVSFEPIHYCAPAGAILKCNKNTFNCTGPTCTNVSTVQCTHGIRPVVSTQLLNG 263  
QY 243 SLAEDEVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTL--GPGRVVYTTGEI 300  
DB 264 SLAEDEVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTL--GPGRVVYTTGEI 323  
QY 301 LGNIROAHNCISRAQWNTLQIATTLREOPG-NKTIAFNOSGGDPEIVMHSFNCGGEF 359  
DB 324 -GNMROAHNCISRAKWSNTLKQIASKLREQFGNKKTIIFKQSSGGDPEIVTHSFNCGGEF 382  
QY 360 FYCNSQLFNSAWNVT-SNGTWSVTRKQDGTGDIITLPCRIKQIINRWVGVKAMVALPI 418  
DB 383 FYCKSTQLFNSAWNVT-SNGTWSVTRKQDGTGDIITLPCRIKQIINRWVGVKAMVALPI 434  
QY 419 KGLIRCSSNITGLLLTRDGGENQTTTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPT 478  
DB 435 SGOIRCSSNITGLLLTRDGGENQTTTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPT 494  
QY 479 KAKRRVQREKRAVGMFLGFLGAAGSTMGATSMALTVOARQLLSGLVQOQNLLRA 538  
DB 495 KAKRRVQREKRAVG-IGALFLGFLGAAGSTMGATSMALTVOARQLLSGLVQOQNLLRA 553  
QY 539 IKAQOHLQLTQVWGIKQOARILAVERYLKQOQLLGFWSGKLICTTAVPWNASWSKT 598  
DB 554 IEAQOHLQLTQVWGIKQOARILAVERYLKQOQLLGFWSGKLICTTAVPWNASWSKT 613  
QY 599 LDQIWNMTWMDREIDNTHYLIYTLIEESQOQEKNOQELLQDKWASL 649  
DB 614 LEQIWNMTWMDREIDNTHYLIYTLIEESQOQEKNOQELLQDKWASL 664

## RESULT 12

US-10-196-515-12  
; Sequence 12, Application US/10196515  
; Publication No. US20030091594A1  
; GENERAL INFORMATION:  
; APPLICANT: HOXIE, James A.  
; APPLICANT: LABRANCHE, Celia C.

; APPLICANT: DOMS, Robert W.  
; APPLICANT: HOFFMAN, Trevor L.  
; TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND  
; TITLE OF INVENTION: THERAPEUTICS  
; FILE REFERENCE: Hoxie 9596-104U1 (0282)  
; CURRENT APPLICATION NUMBER: US/10/196,515  
; CURRENT FILING DATE: 2002-07-16  
; PRIOR APPLICATION NUMBER: US/09/337,387  
; PRIOR FILING DATE: 1999-06-22  
; PRIOR APPLICATION NUMBER: US 09/317,556  
; PRIOR FILING DATE: 1999-05-24  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 759  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-10-196-515-12

Query Match 79.3%; Score 2755; DB 15; Length 759;  
Best Local Similarity 81.3%; Pred. No. 7.8e-241;  
Matches 529; Conservative 43; Mismatches 55; Indels 24; Gaps 9;  
QY 3 LWTVVYGVVWKEATTTILFCASDAKAYDTEAHNVWATHACVPTNPQEVVLENTNF 62  
DB 34 LWTVVYGVVWKEATTTILFCASDAKAYETEVHNVWATHACVPTDNPQEVVLENTNF 93  
QY 63 NWKNNMVEQMHEDIISLWDQSLKPCVKLTPLCVLTNCTDLNNTNTTTELSIIIVWVQ 122  
DB 94 NWKNDMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDL--KNDTNTSSSGRIME- 150  
QY 123 RKGEMRNCSEFNTSIRDKVOREYALFKLDVEPDDNKNTNTNNTKYRLNCTSVITQ 182  
DB 151 --KGEIKNCSEFNTSIRDKVOREYALFKLDVEPDDNKNTNTNNTKYRLNCTSVITQ 203  
QY 183 ACPKVSFEPIHYCTPTGTGTFALLKNDKXFNCTGPTCTNVSTVQCTHGIRPVVSTQLLNG 242  
DB 204 ACPKVSFEPIHYCAPAGAILKCNKNTFNCTGPTCTNVSTVQCTHGIRPVVSTQLLNG 263  
QY 243 SLAEDEVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTL--GPGRVVYTTGEI 300  
DB 264 SLAEDEVIRSENFNTNAKTIIVQLNVSEINCTRPNNHTRKRVTL--GPGRVVYTTGEI 323  
QY 301 LGNIROAHNCISRAQWNTLQIATTLREOPG-NKTIAFNOSGGDPEIVMHSFNCGGEF 359  
DB 324 -GNMROAHNCISRAKWSNTLKQIASKLREQFGNKKTIIFKQSSGGDPEIVTHSFNCGGEF 382  
QY 360 FYCNSQLFNSAWNVT-SNGTWSVTRKQDGTGDIITLPCRIKQIINRWVGVKAMVALPI 418  
DB 383 FYCKSTQLFNSAWNVT-SNGTWSVTRKQDGTGDIITLPCRIKQIINRWVGVKAMVALPI 434  
QY 419 KGLIRCSSNITGLLLTRDGGENQTTTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPT 478  
DB 435 SGOIRCSSNITGLLLTRDGGENQTTTEIFRPGGDMRDNRSELYKYKVVKIEPLGVAPT 494  
QY 479 KAKRRVQREKRAVGMFLGFLGAAGSTMGATSMALTVOARQLLSGLVQOQNLLRA 538  
DB 495 KAKRRVQREKRAVG-IGALFLGFLGAAGSTMGATSMALTVOARQLLSGLVQOQNLLRA 553  
QY 539 IKAQOHLQLTQVWGIKQOARILAVERYLKQOQLLGFWSGKLICTTAVPWNASWSKT 598  
DB 554 IEAQOHLQLTQVWGIKQOARILAVERYLKQOQLLGFWSGKLICTTAVPWNASWSKT 613  
QY 599 LDQIWNMTWMDREIDNTHYLIYTLIEESQOQEKNOQELLQDKWASL 649  
DB 614 LEQIWNMTWMDREIDNTHYLIYTLIEESQOQEKNOQELLQDKWASL 664

## RESULT 13

US-09-891-609-4  
; Sequence 4, Application US/09891609  
; Patent No. US2002012738A1  
; GENERAL INFORMATION:

; APPLICANT: Stamatos, Leonidas  
; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR  
; FILE REFERENCE: 2570-1-002N  
; CURRENT APPLICATION NUMBER: US/09/891,609  
; CURRENT FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/214,608  
; FILING DATE: 2000-06-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 619  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-891-609-4

Query Match 78.5%; Score 2728.5; DB 10; Length 619;  
Best Local Similarity 79.6%; Pred. No. 1.5e-238;  
Matches 520; Conservative 35; Mismatches 39; Indels 59; Gaps 8;  
QY 3 LWTVVYGVVWKEATTTLCASDAKAYDEAHNVAHACVPTNPQEVVLENTNF 62  
Db 6 LWTVVYGVVWKEATTTLCASDAKAYDEAHNVAHACVPTNPQEVVLENTNF 65  
QY 63 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVWEO 122  
Db 66 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTKSN---WKE 119  
QY 123 RKGEMRNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYLINCNSTVITQ 182  
Db 120 MDRGEIKNCSEFKVGAG-----KLINCNTSVITQ 147  
QY 183 ACPKVSFPIPIHYCTPTGPFALLKNDKFKNGTGPCTNVSTVQCTHGRVPVSTOLLNG 242  
Db 148 ACPKVSFPIPIHYCAPAGFAILKNDKFKNGSGPCTNVSTVQCTHGRVPVSTOLLNG 207  
QY 243 SLABEEVVRSENFNTNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVWYTTGEILG 302  
Db 208 SLABEEVVRSENFNTNAKTIIVQLKESVEINCTRPNNHTRKRSITIGPRAFYATGDIIG 267  
QY 303 NIROAHNCISRAQWNTLQIATTLREQFGNKTIAFNOSGGDPRIVMHSFNCGGEPFYC 362  
Db 268 DIROAHNCISGEKWNNTLKQIVTKLQAGFGNKTIVFKOSSGGDPRIVMHSFNCGGEPFYC 327  
QY 363 NSTQIFNSAWNT-----SNGTWSVTRKQKDTGDIITLPCRKQIINRWQVGVKAMYALP 417  
Db 328 NSTQIFNSAWNTIGPNTNGT-----ITLPCRKQIINRWQVGVKAMYAPP 374  
QY 418 IKGLIRCSNITGLLLTRDGGGE-NOTTEIFRPGGDMRDNRSELYKYKVKIEPLGVA 476  
Db 375 IRGQIRCSNITGLLLTRDGGGEISNTTEIFRPGGDMRDNRSELYKYKVKIEPLGVA 434  
QY 477 PTKAKRRVQREKRAVGMFLGAGSTMGATSMALTVOARQLLSGIQQONNLL 536  
Db 435 PTKAKRRVQREKRAV-TLGMFLGAGSTMGARSLTLTVQARQLLSGIQQONNLL 493  
QY 537 RAIKAQOHLQLTWVGIK-LQARVLAVERYLKQOQLGFWCSCGKLICTTAVPNWASWN 596  
Db 494 RAIKAQOHLQLTWVGIK-LQARVLAVERYLKQOQLGFWCSCGKLICTTAVPNWASWN 552  
QY 597 KTLDOINWNTMWEDEIDNTHLYTLIESNQOEKNOQELLQLDKWSL 649  
Db 553 KSLDOIWNNTMWEDEIDNTHLYTLIESNQOEKNOQELLQLDKWSL 605

RESULT 14  
US-09-891-609-2  
; Sequence 2, Application US/09891609  
; Patent No. US20020127238A1  
; GENERAL INFORMATION:  
; APPLICANT: Stamatos, Leonidas  
; TITLE OF INVENTION: HIV-1 VACCINES AND SCREENING METHODS THEREFOR  
; FILE REFERENCE: 2570-1-002N  
; CURRENT APPLICATION NUMBER: US/09/891,609

; CURRENT FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 60/214,608  
; FILING DATE: 2000-06-27  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 646  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus type 1  
US-09-891-609-2

Query Match 78.5%; Score 2728.5; DB 10; Length 646;  
Best Local Similarity 79.6%; Pred. No. 1.6e-238;  
Matches 520; Conservative 35; Mismatches 39; Indels 59; Gaps 8;  
QY 3 LWTVVYGVVWKEATTTLCASDAKAYDEAHNVAHACVPTNPQEVVLENTNF 62  
Db 33 LWTVVYGVVWKEATTTLCASDAKAYDEAHNVAHACVPTNPQEVVLENTNF 92  
QY 63 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTTTELSIIVWEO 122  
Db 93 NMKNMVEQMHEDIISLWDQSLKPCVKLTPLCVTLNCTDLNNTNTKSN---WKE 146  
QY 123 RKGEMRNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYLINCNSTVITQ 182  
Db 147 MDRGEIKNCSEFKVGAG-----KLINCNTSVITQ 174  
QY 183 ACPKVSFPIPIHYCTPTGPFALLKNDKFKNGTGPCTNVSTVQCTHGRVPVSTOLLNG 242  
Db 175 ACPKVSFPIPIHYCAPAGFAILKNDKFKNGSGPCTNVSTVQCTHGRVPVSTOLLNG 234  
QY 243 SLABEEVVRSENFNTNAKTIIVQLNVSVEINCTRPNNHTRKRVTLGPRVWYTTGEILG 302  
Db 235 SLABEEVVRSENFNTNAKTIIVQLKESVEINCTRPNNHTRKRSITIGPRAFYATGDIIG 294  
QY 303 NIROAHNCISRAQWNTLQIATTLREQFGNKTIAFNOSGGDPRIVMHSFNCGGEPFYC 362  
Db 295 DIROAHNCISGEKWNNTLKQIVTKLQAGFGNKTIVFKOSSGGDPRIVMHSFNCGGEPFYC 354  
QY 363 NSTQIFNSAWNT-----SNGTWSVTRKQKDTGDIITLPCRKQIINRWQVGVKAMYALP 417  
Db 355 NSTQIFNSAWNTIGPNTNGT-----ITLPCRKQIINRWQVGVKAMYAPP 401  
QY 418 IKGLIRCSNITGLLLTRDGGGE-NOTTEIFRPGGDMRDNRSELYKYKVKIEPLGVA 476  
Db 402 IRGQIRCSNITGLLLTRDGGGEISNTTEIFRPGGDMRDNRSELYKYKVKIEPLGVA 461  
QY 477 PTKAKRRVQREKRAVGMFLGAGSTMGATSMALTVOARQLLSGIQQONNLL 536  
Db 462 PTKAKRRVQREKRAV-TLGMFLGAGSTMGARSLTLTVQARQLLSGIQQONNLL 520  
QY 537 RAIKAQOHLQLTWVGIK-LQARVLAVERYLKQOQLGFWCSCGKLICTTAVPNWASWN 596  
Db 521 RAIKAQOHLQLTWVGIK-LQARVLAVERYLKQOQLGFWCSCGKLICTTAVPNWASWN 579  
QY 597 KTLDOINWNTMWEDEIDNTHLYTLIESNQOEKNOQELLQLDKWSL 649  
Db 580 KSLDOIWNNTMWEDEIDNTHLYTLIESNQOEKNOQELLQLDKWSL 632

RESULT 15  
US-10-032-162-17  
; Sequence 17, Application US/10032162  
; Publication No. US20030052839A1  
; GENERAL INFORMATION:  
; APPLICANT: BINLEY, JAMES M  
; APPLICANT: SCHUELKE, NORBERT  
; APPLICANT: OLSON, WILLIAM C  
; APPLICANT: PAUL, MADSON J  
; APPLICANT: JOHN, MOORE P  
; TITLE OF INVENTION: STABILIZED VIRAL ENVELOPE PROTEINS AND USES THEREOF  
; FILE REFERENCE: 2048/59331a2  
; CURRENT APPLICATION NUMBER: US/10/032,162

Job time : 39 secs

```
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/602,864
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (151)..(151)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (205)..(205)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (213)..(213)
; OTHER INFORMATION: X=UNKNOWN AMINO ACID
;
US-10-032-162-17

Query Match      78.3%; Score 2720.5; DB 15; Length 625;
Best Local Similarity 80.1%; Pred. No. 7.9e-238;
Matches 518; Conservative 36; Mismatches 56; Indels 37; Gaps 5;

.
QY      3 LWTYYGVVVKKEATTLFCASDAKAYDTAHHVWATHACVPTNPNPQEVWLVNTEVF 62
Db      4 LWTYYGVVVKKEATTLFCASDAKAYDTAHHVWATHACVPTDNPQEVWLVNTEHF 63
QY      63 NMKNMNMVEQMHEDIISLWQSLKPCVKLPLCVTLNCTLNNTNTTNTTSLIIIVWEQ 122
Db      64 NMKNMNMVEQMHEDIISLWQSLKPCVKLPLCVTLNCKDVNATNTTNDSEGM----- 117
QY      123 RGKGMNRCFNITTSIRDVQREYALFYKLDVEPIDDNKNTNTNTKYRLINCNTSVITQ 182
Db      118 -ERGEIKNCNFNITTSIRDVEQYALFYKLDVVDXIDN-----NNTSYRLISCOTSVITQ 171
QY      183 ACPKVSFEPPIHYCTPTGALLKNDKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNG 242
Db      172 ACPKISFEPPIHYCAPAGFAILLKNDKTFNGKXPKNVSTXCTHGIRPVVSTQLLLNG 231
QY      243 SLAEFVVIRSENFNTNAKTIIVQLNVSVETNCTRPNNHTRKRVTLGPRVWYTTGELG 302
Db      232 SLAEFVVIRSDNFTNAKTIIVQLKESVEINCTRPNNNG-----AG 273
QY      303 NIRQAHCNISRAQWNTLQOIATTLREQFGNKTIAPNOSGGDPEIVMHSFNCGGFFYC 362
Db      274 DIRQAHCNISRAKNDTLKQIVIKLREQFNKTIIVNHSGGDPEIVMHSFNCGGFFYC 333
QY      363 NSTQLNSANVTSGTWVTRKQKDTGDIITLPCRKQIINRWQVVGVKAMYALPIKGLI 422
Db      334 NSTQLFNSWNNNTGSG-----NNTGNTITLPCRKQIINRWQEVGVKAMYAPPINGQI 387
QY      423 RCSSNITGLLLTDDGGENOTTIFRPGGDMRDNRSSELYKYKVKVIEPLGVAPTAKR 482
Db      388 RCSSNITGLLLTDDGGINENGTFIFRPGGDMRDNRSSELYKYKVKVIEPLGVAPTCKR 447
QY      483 RVVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVAQRLLSGIVQQQNLLRAIKQ 542
Db      448 RVVQREKRAVG-IGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIVQQQNLLRAIEAQ 506
QY      543 QHLLQLTVWGIGKQIARILAVERYLKDQQLLGFWCGSGKLICTTAVPWNASWNSKTLDOI 602
Db      507 QRMQLTVWGIGKQIARVLAVERYLGDQQLLGIWCGSGKLICTTAVPWNASWNSKSLDRI 566
QY      603 WNNMTWWEWDREIDNTYTHLLIYTLIERSONQOEKNQOELLQDLKWSL 649
Db      567 WNNMTWWEWEREIDNTYTLIERSONQOEKNQOELLQDLKWSL 613
```

Search completed: January 13, 2004, 09:16:34

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 13, 2004, 09:06:27 ; Search time 40 Seconds  
(without alignments)  
4186.903 Million cell updates/sec

Title: US-09-938-406-1\_COPY\_33\_681

Perfect score: 3474

Sequence: 1 ANLWTVYGVFWKREATT.....QNQQKXNQQLLDKWL 649

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.23:\*

1: sp.archaea:\*

2: sp.bacteria:\*

3: sp.fungi:\*

4: sp.human:\*

5: sp.invertebrate:\*

6: sp.mammal:\*

7: sp.mhc:\*

8: sp.organelle:\*

9: sp.phage:\*

10: sp.plant:\*

11: sp.rodent:\*

12: sp.virus:\*

13: sp.vertibrate:\*

14: sp.unclassified:\*

15: sp.virus:\*

16: sp.bacteriap:\*

17: sp.archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2908	83.7	854	15	O40222 human immun
2	2896.5	83.4	853	15	O03811 human immun
3	2889	83.2	841	15	O90KJ1 human immun
4	2872.5	82.7	863	15	O77989 human immun
5	2869.5	82.6	843	15	O70150 human immun
6	2865.5	82.5	851	15	O56110 human immun
7	2865.5	82.5	863	15	O42031 human immun
8	2862.5	82.4	863	15	O9WJU8 human immun
9	2860	82.3	854	15	O9E1S8 human immun
10	2860	82.3	868	15	O9WJW5 human immun
11	2858.5	82.3	863	15	O9WJU4 human immun
12	2857	82.2	858	15	O9DL81 human immun
13	2855.5	82.2	853	15	O9J022 human immun
14	2855.5	82.2	855	15	O9E1R7 human immun
15	2854	82.2	858	15	O9DH32 human immun
16	2853	82.1	864	15	O9YP48 human immun

17	2852	82.1	864	15	O9YP39 human immun
18	2850	82.0	856	15	O9E1R8 human immun
19	2849.5	82.0	847	15	O75760 human immun
20	2846.5	81.9	871	15	O902L4 human immun
21	2845.5	81.9	860	15	O9E1S7 human immun
22	2845	81.9	856	15	O77694 human immun
23	2844.5	81.9	853	15	O9J023 human immun
24	2844.5	81.9	855	15	O8Q367 human immun
25	2844	81.9	856	15	O9E1S5 human immun
26	2844	81.9	856	15	O72993 human immun
27	2843.5	81.9	857	15	O89654 human immun
28	2843.5	81.9	864	15	O9YP31 human immun
29	2842	81.8	854	15	O78225 human immun
30	2841.5	81.8	853	15	O56108 human immun
31	2841	81.8	858	15	O9DL80 human immun
32	2840.5	81.8	849	15	O8Q851 human immun
33	2840.5	81.8	853	15	O9E1S9 human immun
34	2840.5	81.8	859	15	P87924 human immun
35	2840	81.8	858	15	O9DL79 human immun
36	2838	81.7	848	15	O9E1S3 human immun
37	2838	81.7	854	15	O99C08 human immun
38	2838	81.7	860	15	O9YP50 human immun
39	2838	81.7	862	15	O9E1S2 human immun
40	2837	81.7	852	15	O92761 human immun
41	2837	81.7	852	15	O41883 human immun
42	2836.5	81.6	852	15	O73303 human immun
43	2835.5	81.6	851	15	O8Q852 human immun
44	2835.5	81.6	861	15	O9WJV5 human immun
45	2835	81.6	735	15	O9IW16 human immun

ALIGNMENTS

RESULT 1

O40222	PRELIMINARY;	PRT;	854 AA.
ID	O40222		
AC	O40222;		
DT	01-JAN-1998 (Tremblrel. 05, Created)		
DT	01-JAN-1998 (Tremblrel. 05, Last sequence update)		
DT	01-OCT-2002 (Tremblrel. 22, Last annotation update)		
DE	Env polypeptide.		
GN	ENV.		
OS	Human immunodeficiency virus 1.		
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.		
OX	NCBI_TaxID=11676;		
RN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=AD8;		
RX	MEDLINE=96432129; PubMed=8835195;		
RA	Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,		
RA	Martin M.A., Peden K.W.;		
RT	"Construction and characterization of a stable full-length macrophage-		
RT	tropic HIV type 1 molecular clone that directs the production of high		
RT	titers of progeny virions.";		
RL	Aids Res. Hum. Retroviruses 12:191-194(1996).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=AD8;		
RA	Theodore T.S., Englund G., Buckler-White A., Buckler C.E.,		
RA	Martin M.A., Peden K.W.C.;		
RL	Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF004394; AAB64170.1;		
DR	InterPro; IPR00328; Env GP41.		
DR	InterPro; IPR00777; GP120.		
DR	Pfam; PF00516; GP120; 1.		
DR	Pfam; PF00517; GP41; 1.		
KW	SEDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.		
SQ	SEQUENCE 854 AA; 972391 MW; 06C45E69103C6C12 CRC64;		

Query Match 83.7%; Score 2908; DB 15; Length 854;  
Best Local Similarity 84.0%; Pred. No. 1.3e-230;  
Matches 548; Conservative 36; Mismatches 48; Indels 20; Gaps 7;



DR	EMBL; AF112547; AAF13325.1; -.
DR	InterPro; IPR000328; Env_GP41.
DR	InterPro; IPR000777; GP120.
DR	Pfam; PF00516; GP120; 1.
DR	Pfam; PF00517; GP41; 1.
KW	AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT	NON_TER
SQ	SEQUENCE 841 AA; 95389 MW; E9086659E7E1D447 CRC64;
	Query Match 83.2%; Score 2889; DB 15; Length 841;
	Best Local Similarity 83.5%; Pred. No. 4.6e-229;
	Matches 542; Conservative 38; Mismatches 53; Indels 16; Gaps 4;
QY	3 LWTYYGVVWK <sup>1</sup> EATTTLCASDAKAYDEAHNVWATHACVPINPQEVLENVTENF 62
DG	
DB	20 LWTVYYGVVWK <sup>1</sup> EATTTLCASDAKAYDEVHNVWATHACVPIDPSQEVLENVTENF 79
QY	63 NMKNNNVEQM <sup>HED</sup> IISLWDSLKPCVKLTPLCVTLNCTDLNTNTTNTTSLSIIVWEQ 122
DB	
DB	80 NMKNNNVEQM <sup>HED</sup> IISLWDQS <sup>LKPCVKLTPLCVTLNCTDNALNATKTNS</sup> NSSL---- 135
QY	123 RGKGEMNCNFNI <sup>TTS</sup> IRDKVOREYALFYKLDDVIPDDNKNTNNTKYRLINCNTSVITQ 182
DB	
DB	136 EMKGMCNCFNI <sup>TTS</sup> IRDKVOREYALFYKLDDVIPGSD---SNTSYRLINCNTSVITQ 192
QY	183 ACPKVSPEPIPHYCTPTGFALLKCNDKKFNGTGPCTNVSTVOCTHGIRPVVSTQLLNG 242
DB	
DB	193 ACPKVSPEPIPHCCAPAGFAILKCNDKKFNGTGPCTNVSTVOCTHGIPKVVSTQLLNG 252
QY	243 SLAEEVVIRSENFTNNAKI <sup>II</sup> VOLNVSEINCTRPNNHTRKRVTLGPGRVWYTTEILG 302
DB	
DB	253 SLAEEVVIRSENFTDNAKI <sup>II</sup> VHLNESVITCTRPNNNTRKG:HIGPRTFYTTGEIIG 312
QY	303 NIRAQHCHISRAOWNNTLOLIATTLRQFGNKTI <sup>IAFNOSSGDDEI</sup> VMHSFCGGEFFYC 362
DB	
DB	313 DIRAQHNLSRTQNNTLRQVKKLRQFGNKTI <sup>IVFNQSSGGDDEI</sup> MHSFCGGEFFYC 372
QY	363 NSTOLFNSANWITSNGVSWSTRKOKDTGDITTLPCR <sup>IKQIINRWQVVGKAM</sup> YALPIKGLI 422
DB	
DB	373 NSTOLFNSWNVTESNSSI-----GENITLPCR <sup>IKQIVNMQKVGKAMYAP</sup> PIRQI 425
QY	423 RCSSNITGLLITRDGGENGQT--TEIFRPGGDMRDNRSEL <sup>YKYKVKVKB</sup> PLGVAPTKA 480
DB	
DB	426 RCTSNITGLLLTRDGGNNNSNETETFRPGGDMRDNRSEL <sup>YKYKVKVKB</sup> PLGVAPTKA 485
QY	481 KRRVVQREKAVGM <sup>LGAFLGCAAGSTM</sup> CATSMALTVAQRLLSGIVQQNNLLRAIK 540
DB	
DB	486 KRRVVQREKAVGM <sup>LGAFLGCAAGSTM</sup> AAISITLVQAQRLLSGIVQQNNLLRAIE 545
QY	541 AQOHLQLTYWGIRKQLOARILAVERYLKDQOLLGFWGCGSKLI <sup>CTTAVPMNAS</sup> NSKNTLD 600
DB	
DB	546 AQOHLQLTYWGIRKQLOARILAVERYLKDQOLLGFWGCGSKLI <sup>CTTIVPMN</sup> SSWSKSLD 605
QY	601 QIWNMTWMERIDN <sup>IDNYTHLYTLIESQ</sup> NQOEKQOELLQLODKWASL 649
DB	
DB	606 KIWNMTWMERIDN <sup>YTSIYTLLEV</sup> SQNEKELELDELDKWASL 654
RESULT 4	
Q77989	PRELIMINARY; PRT; 863 AA.
ID	Q77989
AC	Q77989;
DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE	Envelope glycoprotein.
GN	ENV.
OS	Human immunodeficiency virus 1.
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX	NCBI_TaxID=11676;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=SFI28A;

RX	MEDLINE=91056585; PubMed=2243391;
RA	Liu Z.-Q., Wood C., Levy J.A., Cheng-Mayer C.;
RT	"The viral envelope gene is involved in macrophage tropism of a human
RL	immunodeficiency virus type 1 strain isolated from brain tissue.";
RL	J. Virol. 64:6148-6153 (1990).
DR	EMBL; M95292; AA844331.1; -:-
DR	InterPro; IPR000328; Env_GP41.
DR	InterPro; IPR000777; GP120.
DR	Pfam; PF00516; GP120; 1.
DR	Pfam; PF00517; GP41; 1.
KW	AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ	SEQUENCE 863 AA; 98222 NW; 0157397FA8BAC5D7 CRC64;
Query Match 82.7%; Score 2872.5; DB 15; Length 863;	
Best Local Similarity 82.6%; Pred. No. 1.le-227;	
Matches 543; Conservative 43; Mismatches 48; Indels 23; Gaps	
QY	3 LMWTVVYGVPWKAEATTTTLFCASDAKAYDTAAHNVWATHACVTPTRNPNPOEVLVENTENF 62
Dd	33 LMWTVVYGVPWKAEATTTTLFCASDAKADSTEVHNVWATHACVPTDPNPQPEVLVENTENF 92
QY	63 NMWKNMWQEHEDIISLWDQSLPCVKLTPLCVLNCT-----DLNTNNNTTNTTS 115
Dd	93 NMWKNMWQEHEDIISLWDQSLPCVKLTPLCVLNCTDNLDNDNTND-TNATSSS 151
QY	116 IIVWEQQRK-GEMNCSEFNITTSIRDKVQREYAFYKL DVEPIDDKNTTKRYRLIN 174
Dd	152 L-----RGETGEIKNCSFNITTSIRDKVQREYAFYKL DIVPINN----NNTTYRLIN 200
QY	175 CNTSVITQACP KVSFEPIPHYCHTPTGFALLKCKNDKFNGTGCTNVTSTVOCTHGIRPVV 234
Dd	201 CNTSVITQACP KVSFEPIPHYCHT PAGFAILKCKNDKFNGKGPCKNVSTVOCTHGIRPVV 260
QY	235 STOLLINGS LAEEVVIRESEFTNNAKTIIIVOLNVSYEINCTRPNNHTKRVTGLGPRVM 294
Dd	261 STOLLINGS LAEEVVIRESDNFNTNAKTIIIVOLNESVINCTRPNNTRKSINIGFORAI 320
QY	295 YTTGTILGNIROAHCNISRAQWNNTLQQIATTLRREQFGNKTI AFNQSGGDPEIVMHSEFN 354
Dd	321 YTTGAIIIGDIROAHCTLNKQTQWDNLTRIAIKLRQEFKNKTI SFKQSSGGDPEIVMQSFN 380
QY	355 CGGEFFYCYNSTOLFNSANWVTS--NGWTSVTRKOKTDGITILPCR IKQLINRWQVVGKA 412
Dd	381 CGGEFFYCYN TTKLFNSTWDRFLFNSTWNNTIGSNTEGNIITLP CRIKQLINMWQEVGKA 440
QY	413 MYALPTKGLIRCSSNITGLLTRDGGGENOTETFRPGGDMRDNRSELYKYKWKIEP 472
Dd	441 MTAPPIRGQICSSNITGLLTRDGGNETETEFRPGGDMRDNRSELYKYKWKIEP 500
QY	473 LGVAPTAKRRVVQREKRAVGMGLGAMFLGFLGAAGSTMWATSMA LTVQARQLLSGVVQQ 532
Dd	501 LGVAPTAKRRVVQREKRAVG-IGAVFLGFLGAAGTMWAASVAL TVQARQLLSGVVQQ 559
QY	533 NLLRAIKAQHLLQITVWGVIKQIQARILAVERYLKQQLLGFWCGSGSKLIC TTAPFWNA 592
Dd	560 NLLRAIEAQHLLQITVWGVIKQIQARVLAVERYLRDQQLLGITGWCGSGSKLIC TTVPWNA 619
QY	593 SWSNKTLDQIWNMTWMSEIDRNTHLIYTLTEESQNQEKKQQLQLD KWASL 649
Dd	620 SWSNKSLDKWNMNTWMSEWEREIDNYTSLIYTLTEESQNQEKKQEQLLEDD KWASL 676
RESULT 5	
Q70150	
ID	Q70150 PRELIMINARY; PRT; 843 AA.
AC	Q70150;
DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE	Envelope glycoprotein (Fragment).
GN	ENV.
OS	Human immunodeficiency virus 1.
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.

```

OX NCBI_TaxID=11676;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=014;
RX MEDLINE=95194694; PubMed=7888189;
RA Gao F., Yue L., Craig S., Thornton C.L., Robertson D.L.,
RA McCutchan F.E., Bradac J.A., Sharp P.M., Hahn B.H.;
RA "Genetic variation of HIV type 1 in four World Health Organization-
RT sponsored vaccine evaluation sites: generation of functional envelope
RT (glycoprotein 160) clones representative of sequence subtypes A, B, C,
RT and E. WHO Network for HIV Isolation and Characterization.";
RL AIDS Res. Hum. Retroviruses 10:1359-1368(1994).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=014;
RX MEDLINE=95194691; PubMed=7545977;
RA WHO Global Programme on AIDS;
RT "HIV type 1 variation in World Health Organization-sponsored vaccine
RT evaluation sites: genetic screening, sequence analysis, and
RT preliminary biological characterization of selected viral strains. WHO
RT Network for HIV Isolation and Characterization.";
RL AIDS Res. Hum. Retroviruses 10:1327-1343(1994).
RN [3];
RP SEQUENCE FROM N.A.
RC STRAIN=014;
RX MEDLINE=96190564; PubMed=8627686;
RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,
RA Karlsson G., Sodroski J., Morgado M., Galvao-Castro B.,
RA von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,
RA Hahn B.H.;
RA "Molecular cloning and analysis of functional envelope genes from
RT human immunodeficiency virus type 1 sequence subtypes A through G. The
RT WHO and NIAID Networks for HIV Isolation and Characterization.";
RL J. Virol. 70:1651-1657(1996).
RN [4];
RP SEQUENCE FROM N.A.
RC STRAIN=014;
RX Allen E.E.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U08801; AAB05185.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON TER
SQ SEQUENCE 843 AA; 95679 MW; 14DB49A6D2F5FCE4 CRC64;

Query Match 82.6%; Score 2869.5; DB 15; Length 843;
Best Local Similarity 83.4%; Pred. No. 1.8e-227;
Matches 542; Conservative 39; Mismatches 42; Indels 27; Gaps 8;

QY 3 LWTVVYGVVWKEATTTLCASDAKAYDTBAHNVWATHACVPTNPQEVVLENTENF 62
DB 31 LWTVVYGVVWKEATTTLCASDAKAYDTBAHNVWATHACVPTNPQEVVLENTENF 90
QY 63 NWNKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDL-NTNNTNTTSLIIVWE 121
DB 91 NWNKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDL-NTNNTNTTSLIIVWE 142
QY 122 QRGKEMENCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCNTSVIT 181
DB 143 ---GGEIKNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCNTSVIT 193
QY 182 QACPKVSEPIPIHYCTPTGTFALLKCNCKKNGTGPCNTNSTVQCTHGRVWSTQLLN 241
DB 194 QACPKVSEPIPIHYCTPTGTFALLKCNCKKNGTGPCNTNSTVQCTHGRVWSTQLLN 253
QY 242 GSLAEVEEVVISENTNNAKTIIVQLNVSVEINCTRNPNHTRKRVTLGPGRVWVTGAIL 301
DB 254 GSLAEVEEVVISENTNNAKTIIVQLNVSVEINCTRNPNHTRKRVTLGPGRVWVTGAIL 313
QY 302 GNIRQAHCNISRAQWNNTLQIATTLREQFGNKTIAFNQSSGGDPEIWMHSFNCGGEFFY 361

```

```

DB 314 GDIRQAHCNLSSTKWNNTLRQITEKLRQFGNKTIIVNQSSGGDPEIWMHSFNCGGEFFY 373
QY 362 CNSTQLFNSAWNTSNGTWSVTRKQKDTGD-IITLPCRIKQIINRWQVVGKAMALPKG 420
DB 374 CNTQLFNSWNTDTS--TWN-----NNTNGTITLPCRIKQIINRWQVVGKAMALPKG 426
QY 421 LIRCSSNITGLLLTRDGG-GENOTTEIFRPGGGMDRNRSELYKYKVKIEPLGVAPTK 479
DB 427 QIRCSSNITGLLLTRDGGSENKITEIFRPGGGMDRNRSELYKYKVKIEPLGVAPTK 486
QY 480 AKREVQREKRAVGMGLGAMFLGAGSTMGATSMALTVOARQLLSGIVQQNNLLRAI 539
DB 487 PKREVQREKRAVGTIGAMFLGAGSTMGATSMALTVOARQLLSGIVQQNNLLRAI 546
QY 540 KAOQHLLQLTWGIGKQIQAIRLAVERLYKQQLLGFWGCSGLICTTAVPWNASWNTKL 599
DB 547 EAOQHLLQLTWGIGKQIQAIRLAVERLYKQQLLGFWGCSGLICTTAVPWNASWNTKL 606
QY 600 DOIWNNTWMDREIDNYTHLYTLIEESQOQEKQOELLQDKWASL 649
DB 607 DKIWNTWMDREIDNYTHLYTLIEESQOQEKQOELLQDKWASL 656

RESULT 6
O56110 PRELIMINARY; PRT; 851 AA.
AC O56110;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=SFVHS8.1;
RX MEDLINE=9818716; PubMed=9519894;
RA McCutchan F.E., Sanders-Buell E., Salminen M.O., Carr J.K.,
RA Sheppard W.H.;
RA "Diversity of the human immunodeficiency virus type 1 envelope
RT glycoprotein in San Francisco Men's Health Study participants.";
RL AIDS Res. Hum. Retroviruses 14:329-337(1998).
DR EMBL; AF025756; AAC40593.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 851 AA; 96516 MW; 1CDAC83BE6464531 CRC64;

Query Match 82.5%; Score 2865.5; DB 15; Length 851;
Best Local Similarity 83.2%; Pred. No. 4e-227;
Matches 538; Conservative 42; Mismatches 52; Indels 15; Gaps 5;

QY 3 LWTVVYGVVWKEATTTLCASDAKAYDTBAHNVWATHACVPTNPQEVVLENTENF 62
DB 33 LWTVVYGVVWKEATTTLCASDAKAYDTBAHNVWATHACVPTNPQEVVLENTENF 92
QY 63 NWNKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDL-NTNNTNTTSLIIVWEQ 122
DB 93 NWNKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDL-NTNNTNTTSLIIVWEQ 146
QY 123 RKGEMENCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCNTSVITQ 182
DB 147 MERGEIKNCSEFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNTKYRLINCNTSVITQ 200
QY 183 ACPKVSSEPIPIHYCTPTGTFALLKCNCKKNGTGPCNTNSTVQCTHGRVWSTQLLN 242

```



Db 201 ACPKVSFPIPIHYCAPAGFAILKCNDRKPNKPGCTNVSTVQCTHGIRPVVSTQPLNG 260  
 QY 243 SLAEVWIRSENFNTNNAKTIIVQLNVSVINCTRPNNHTRKRVTLGPRGVWYTTGEILG 302  
 Db 261 SLAEVWIRSENFNTNNAKTIIVQLNVSVINCTRPNNHTRKRVTLGPRGVWYTTGEILG 320  
 QY 303 NTRQAHNISRAQWNTLQOIAATTIAREQPKNTIAPNOSGGDPPIVHMSFNCGGEFFYC 362  
 Db 321 DIRQAHNLSEAKNHTLEQIAKLRQEGNKTIIVFNOSGGDPPIVHMSFNCGGEFFYC 380  
 QY 363 NSTQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAMVALPIKGLI 422  
 Db 381 NSTKLFNRTSV--NSTWMDTEGVNTGENITLPCRIKQIINRWQVGVKAMVALPIKGLI 438  
 QY 423 RCSSNITGLLTRDGGNGNQTETPRPGGDMRDNRSELYKYKVKIPLGVAPTAKR 482  
 Db 439 RCSSNITGLLTRDGGNSNETETPRPGGDMRDNRSELYKYKVKIPLGVAPTAKR 498  
 QY 493 RVVQREKRAVGMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRAIKAO 542  
 Db 499 RVVQREKRAVG-IGAVFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQNLLRAIEAQ 557  
 QY 543 QHLLQLTVMGIKQOARILAVERYLKQQLIGFWGSGKLICTTAVPWNASNSKTLDOI 602  
 Db 558 QHLLQLTVMGIKQOARILAVERYLKQQLIGFWGSGKLICTTAVPWNASNSKTLDOI 617  
 QY 603 MNMTWMEDREIDNYTHLYTLIESQNKQOELLQDKWASL 649  
 Db 618 MNMTWMEDREIDNYTHLYTLIESQNKQOELLQDKWASL 664

RESULT 7  
 Q42031  
 ID O42031 PRELIMINARY; PRT; 863 AA.  
 AC O42031;  
 DT 01-JAN-1998 (TremBLrel. 05, Created)  
 DT 01-JAN-1998 (TremBLrel. 05, Last sequence update)  
 DT 01-OCT-2002 (TremBLrel. 22, Last annotation update)  
 DE Envelope polyprotein.  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ENL4-3;  
 RA Fang G., Weiser B., Visosky A., Burger H.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF003888; AAB64286.1; -;  
 DR EMBL; AF003887; AAB64278.1; -;  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 KW Transmembrane.  
 SQ SEQUENCE 863 AA; 98166 MW; 918DD40F4F0C16E7 CRC64;

Query Match 82.5%; Score 2865.5; DB 15; Length 863;  
 Best Local Similarity 82.3%; Pred. No. 4.1e-227;  
 Matches 540; Conservative 40; Mismatches 53; Indels 23; Gaps 5;

QY 4 WTVVYGVVPWKEATTLFCASDAKAYDEAHNVWATHACVPTNPNQEVLENVTENFN 63  
 Db 34 WTVVYGVVPWKEATTLFCASDAKAYDEAHNVWATHACVPTNPNQEVLENVTENFN 93  
 QY 64 MWKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTD-----LNTNNTNITELSI 116  
 Db 94 MWKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDNLKNATVNNANNTNNS- 149  
 QY 117 IVVWEORGKGMENCSFNITTSIRDKVQREYALFYKLDVEPID--DNKNTTNTKYLRLN 174  
 Db 150 ---WEKWEKGEIKNCSEFNITTSIRDKVQREYALFYKLDVVPIDNANNATNTYTSRLIS 206

QY 175 CNTSVITQACPKVSFPIPIHYCTPTGFALLKCNDRKPNKPGCTNVSTVQCTHGIRPVV 234  
 Db 207 CNTSVITQACPKVSFPIPIHYCAPAGFAILKCNDRKPNKPGCTNVSTVQCTHGIRPVV 266  
 QY 235 STOLLNGLSLAEVWIRSENFNTNNAKTIIVQLNVSVINCTRPNNHTRKRVTLGPRGVW 294  
 Db 267 STOLLNGLSLAEVWIRSENFNTNNAKTIIVQLNVSVINCTRPNNHTRKRVTLGPRGVW 326  
 QY 295 YTTGILNIRQAHCNISRAQWNTLQOIAATTIAREQPKNTIAPNOSGGDPPIVHMSF 353  
 Db 327 YTTGDIIGDIRQAHCNVSRAKNNTLAVIKLEQFGHNKTIIVFNHSSGGDLIVTHSF 386  
 QY 354 NCGGFFVYCNSTOLPNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAM 413  
 Db 387 ICGGFFVYCNSTOLPNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVGVKAM 440  
 QY 414 VALPIKGLIRCSSNITGLLTRDGGNGNQTETPRPGGDMRDNRSELYKYKVKIPL 473  
 Db 441 YAPPTRGQIRCLSNITGLLTRDGGDTNGTEVFRPGGDMRDNRSELYKYKVKIPL 500  
 QY 474 GVAPTAKRRVVRQREKRAVGMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQN 533  
 Db 501 GVAPTAKRRVVRQREKRAVGMFLGFLGAAGSTMGASMTLTVQARLLSGIVQOQN 560  
 QY 534 NLLRAIKAQHLLQLTVMGIKQOARILAVERYLKQQLIGFWGSGKLICTTAVPWNAS 593  
 Db 561 NLLRAIEAQHLLQLTVMGIKQOARILAVERYLKQQLIGFWGSGKLICTTAVPWNAS 620  
 QY 594 WSKNTLDOIWNMTWMEDREIDNYTHLYTLIESQNKQOELLQDKWASL 649  
 Db 621 WSKNSLNIWNMTWMEDREIDNYTHLYTLIESQNKQOELLQDKWASL 676

RESULT 8  
 Q9WJ08  
 ID Q9WJ08 PRELIMINARY; PRT; 863 AA.  
 AC Q9WJ08;  
 DT 01-NOV-1999 (TremBLrel. 12, Created)  
 DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)  
 DT 01-OCT-2002 (TremBLrel. 22, Last annotation update)  
 DE Envelope polyprotein.  
 GN ENV.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fang G., Weiser B., Chappey C., Visosky A., Townsend L., Wang Q.,  
 RA Burger H.;  
 RT "Complete plasma HIV-1 Sequence: Spl-Promoter Deletion Can Lead to  
 RT Non-Progressive Infection."  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U69590; AAD10923.1; -;  
 DR InterPro; IPR000328; Env GP41.  
 DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SQ SEQUENCE 863 AA; 98148 MW; 5275ACD8476BAE02 CRC64;

Query Match 82.4%; Score 2862.5; DB 15; Length 863;  
 Best Local Similarity 82.3%; Pred. No. 7.2e-227;  
 Matches 540; Conservative 39; Mismatches 54; Indels 23; Gaps 5;

QY 4 WTVVYGVVPWKEATTLFCASDAKAYDEAHNVWATHACVPTNPNQEVLENVTENFN 63  
 Db 34 WTVVYGVVPWKEATTLFCASDAKAYDEAHNVWATHACVPTNPNQEVLENVTENFN 93  
 QY 64 MWKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTD-----LNTNNTNITELSI 116  
 Db 94 MWKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCTDNLKNATVNNANNTNNS- 149

```

QY 117 IUVWEQKGBMRNCSENIITTSIRDKVQREYALFYKLDVEPID--DNKNTTNTTKYRLIN 174
Db 150 ---WEKMEKGEIKKCNFNITTSIRDKVQREYALFYKLDVDPIDRANNSNATNTSYRLIS 206
QY 175 CNTSVITQACPKVSFEPIPIHYCTPTGFALLKCNCKKFNKGPGCTNVSTVQCTHGIRPV 234
Db 207 CNTSVITQACPKVSFEPIPIHYCAPAGFALLKCNCKKFNKGPGCTNVSTVQCTHGIRPV 266
QY 235 STQLLNGSLAEBEVEVIRSENFNTNNAKIIIVQLNVSVIEINCTRNHNHTRKRVTTGPGRVW 294
Db 267 STQLLNGSLAEBEVEVIRSENFNTNNAKIIIVQLNVSVIEINCTRNHNHTRKRVTTGPGRVW 326
QY 295 YTTGEILGNIRQAHCHNISRAQWNTLQOIATTLRBPFG-NKTIAFNOSGGDPEIVMHSF 353
Db 327 YTTGDIIGDIRQAHCHNVSRAKWNNTLVKIVKLEQFGHNKTIIVFNHSSGGDLEIVHSP 386
QY 354 NCGGEFFYCNSTQLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRKIQIINRWQVVGKAM 413
Db 387 ICGGEFFYCNSTQLF-----TNSTWNNTRESNDNTEIILPCRKIQIINRWQVVGKAM 440
QY 414 YALPIKGLIRCSSNITGLLLTRDGGENQTTETFRPGGDMRDNRSELKYKYVVKIEPL 473
Db 441 YAPPIRGQIRCLSNITGLLLTRDGGDTENGTEVFRPGGDMRDNRSELKYKYVVKIEPL 500
QY 474 GVAPTAKRRVVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOON 533
Db 501 GVAPTAKRRVVQREKRAVGTIGAMFLGFLGAAGSTMGAASTLTVOARQLLSGIVQOON 560
QY 534 NLLRAIKAAQHLQLTWGKIQOLQARILAVERYLKDQQLLFGWCSGKLICTTAVPNWAS 593
Db 561 NLLRAIEAQHLQLTWGKIQOLQARVLAVERYLKDQQLLFGWCSGKLICTTAVPNWAS 620
QY 594 WSKNTLQIWNNTWMEWDREIDNTHLYTLIBESQKQKQELLOLDKWSL 649
Db 621 WSKSLNEIWNNTWMEWEREINNYTDLIYTLIBESQKQKQELLELDKWSL 676

RESULT 9
Q9E1S8 PRELIMINARY; PRT; 854 AA.
AC Q9E1S8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope protein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20442410; PubMed=10984542;
RA Cleghorn F.R., Jack N., Carr J.K., Edwards J., Mahabir B., Sill A.,
RA McDaniel C.B., Connolly S.M., Goodman D., Bennetts R.Q., O'Brien T.R.,
RA Weinhold K.J., Bartholomew C., Blattner W.A., Greenberg M.L.;
RT "A distinctive clade B HIV type 1 is heterosexually transmitted in
RT Trinidad and Tobago.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:10532-10537(2000).
DR EMBL; AF277057; AAG22502.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 854 AA; 97210 MW; 2DBA931678837491 CRC64;

Query Match 82.3%; Score 2860; DB 15; Length 854;
Best Local Similarity 82.9%; Pred. No. 1.1e-226;
Matches 535; Conservative 52; Mismatches 46; Indels 20; Gaps 7;

QY 1 ANLWTVYGVVPMKEATTTLCASDAKAYDTEAHNVWATHACVPTDNPQEVVLENVTE 60
Db 31 ANLWTVYGVVPMKEATTTLCASDAKAYDTEAHNVWATHACVPTDNPQEVVLENVTE 90

```

```

QY 61 NFNMKNMVEQMEHDIISLWDSIKPCVKLTPLCVTLNCTDLNTNT-TNTTSLIIV 119
Db 91 NFNMKNMVEQMEHDIISLWDSIKPCVKLTPLCVTLNCTDVRNNSGVNITS----- 144
QY 120 WEQKGBMRNCSENIITTSIRDKVQREYALFYKLDVEPIDDN--KNNTNTKYRLINCN 177
Db 145 ---GKEIKNCSNITTAVERDKVQYALFYRLDVPIDDDHGNSSSNKYRLINCN 200
QY 178 SVITQACPKVSFEPIPIHYCTPTGFALLKCNCKKFNKGPGCTNVSTVQCTHGIRPVSTQ 237
Db 201 SVITQACPKVSFEPIPIHYCAPAGFALLKCNCKKFNKGPGCTNVSTVQCTHGIRPVSTQ 260
QY 238 LLNGSLAEBEVEVIRSENFNTNNAKIIIVQLNVSVIEINCTRNHNHTRKRVTTGPGRVW 297
Db 261 LLNGSLAEBEVEVIRSENFNTNNAKIIIVQLNVSVIEINCTRNHNHTRKRVTTGPGRVW 319
QY 298 GEILGNIRQAHCHNISRAQWNTLQOIATTLRBPFG-NKTIAFNOSGGDPEIVMHSF 357
Db 320 GEILGNIRQAHCHNISRAQWNTLQOIATTLRBPFG-NKTIAFNOSGGDPEIVMHSF 379
QY 358 EFFYCNSTQLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRKIQIINRWQVVGKAM 417
Db 380 EFFYCNSTQLFNSAMNVTSGTWSVTRKQKDTGDIITLPCRKIQIINRWQVVGKAM 434
QY 418 IKGLIRCSSNITGLLLTRDGGENQTTETFRPGGDMRDNRSELKYKYVVKIEPL 476
Db 435 IRGQIRCSSNITGLLLTRDGGENQTTETFRPGGDMRDNRSELKYKYVVKIEPL 494
QY 477 PTAKRRVVQREKRAVGMGLGAMFLGFLGAAGSTMGATSMALTVOARQLLSGIVQOON 536
Db 495 PTAKRRVVQREKRAVGTIGAMFLGFLGAAGSTMGAASTLTVOARQLLSGIVQOON 554
QY 537 RAIKAAQHLQLTWGKIQOLQARILAVERYLKDQQLLFGWCSGKLICTTAVPNWAS 596
Db 555 RAIKAAQHLQLTWGKIQOLQARVLAVERYLKDQQLLFGWCSGKLICTTAVPNWAS 614
QY 597 KTLQIWNNTWMEWDREIDNTHLYTLIBESQKQKQELLOLDKWSL 649
Db 615 KSLSDIWNNTWMEWDREIDNTHLYTLIBESQKQKQELLELDKWSL 667

RESULT 10
Q9WJW5 PRELIMINARY; PRT; 868 AA.
AC Q9WJW5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope polyprotein.
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Fang G., Weiser B., Chappey C., Visosky A., Townsend L., Wang Q.,
RA Burger H.;
RT "Complete Plasma HIV-1 Sequence: Spl-Promoter Deletion Can Lead to
RT Non-Progressive Infection.";
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U69584; AAD10875.1; -
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 868 AA; 98569 MW; F2D138F0F8BFFC02 CRC64;

Query Match 82.3%; Score 2860; DB 15; Length 868;
Best Local Similarity 82.6%; Pred. No. 1.2e-226;
Matches 543; Conservative 36; Mismatches 60; Indels 18; Gaps 7;

```

QY 3 LWVTYVYGVPMKEATTTILFCASDAKAYDEAHNVWATHACVPTNPNQEVLENTENF 62  
DB 33 LWVTYVYGVPMKEATTTILFCASDAKAYDEAHNVWATHACVPTNPNQEVLENTENF 92  
QY 63 NWNKNNVQEMHEDIISLWDQSLKPCVKLTPLCVTLNCTD---LNTNNTNTTETLSIIV 118  
DB 93 NWNKNNVQEMHEDIISLWDQSLKPCVKLTPLCVTLNCTD---LNTNNTNTTETLSIIV 149  
QY 119 VMEQRKGMKNCSEFNITTSIRDKVQREVALFYKLDVEPID--DNKN-TTNNTKYELINC 175  
DB 150 -WGMERGEIKNGSEFNITTSIRDKVQREVALFYKLDVEPID--DNKN-TTNNTKYELINC 208  
QY 176 NTSVITQACPKVSFEPIPIHYCTPTGFAILLKNDKFKNGTGPCTNVSTVQCTHGIRPV 325  
DB 209 NTSVITQACPKVSFEPIPIHYCAPAGFAILLKNDKFKNGTGPCTNVSTVQCTHGIRPV 368  
QY 236 TQLLNGSLAEVEVIRSENFNTNNAKTIIVQLNVSEINCTRENNTKRKRVTLGPGRVMY 295  
DB 269 TQLLNGSLAEVEVIRSENFNTNNAKTIIVQLNVSEINCTRENNTKRKRVTLGPGRVMY 328  
QY 296 TTGEILGNTRQAHNCISRAQWNNTLOQIATTLREQFNGK-TIAPNOSGGDPPIVMSHFN 354  
DB 329 TTGEIIGDIRQAHCTLNREWNNTLAKITEKLEQFNGNITIVFNHSSGGDPPIVMSHFI 388  
QY 355 CGGEFFCYNSTOLFNSAWNTSGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKAY 414  
DB 389 CGGEFFCYNSTOLFNSAWNTSGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGKAY 444  
QY 415 ALPIKILRCSSNITGLLLTRDGGG--ENQTTIFRPGGDMRDNRSELYKYKVKVIEP 472  
DB 445 APPIRGQIRCSNITGLLLTRDGGGNTDENRTHIFRPGGDMRDNRSELYKYKVKVIEP 504  
QY 473 LGVAPTKARRVVQREKRAVGMFLGAGTGMGATSMALTVQARQLLSGIVQOQ 532  
DB 505 LGVAPTKARRVVQREKRAVGMFLGAGTGMGATSMALTVQARQLLSGIVQOQ 564  
QY 533 NLLRAIKAQHLLQITVWGIKQARILAVERYLKDQQLGFWGCGKLICTTAVPWN 592  
DB 565 NLLRAIKAQHLLQITVWGIKQARILAVERYLKDQQLGFWGCGKLICTTAVPWN 624  
QY 593 SWSNKTLDQIWNMTWMEWDREIDNTHLYTLIESQKQKNOQELLQDKWASL 649  
DB 625 SWSNKSLEINWMTWMEWDREIDNTHLYTLIESQKQKNOQELLQDKWASL 681

## RESULT 11

Q9WJU4 PRELIMINARY; PRT; 863 AA.  
ID Q9WJU4  
AC Q9WJU4  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
EN Envelope polyprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fang G., Weiser B., Chappey C., Visosky A., Townsend L., Wang Q.,  
RA Burger H.,  
RT "Complete Plasma HIV-1 Sequence: Spl-Promoter Deletion Can Lead to  
RT Non-Progressive Infection."  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U69592; AAD10939.1; -  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Transmembrane.  
SQ SEQUENCE 863 AA; 98269 MW; BE12C858E9CE4D11 CRC64;

Query Match

82.3%; Score 2858.5; DB 15; Length 863;

Best Local Similarity 81.7%; Pred. No. 1.5e-226;  
Matches 537; Conservative 44; Mismatches 53; Indels 23; Gaps 5;  
QY 3 LWVTYVYGVPMKEATTTILFCASDAKAYDEAHNVWATHACVPTNPNQEVLENTENF 62  
DB 33 LWVTYVYGVPMKEATTTILFCASDAKAYDEAHNVWATHACVPTNPNQEVLENTENF 92  
QY 63 NWNKNNVQEMHEDIISLWDQSLKPCVKLTPLCVTLNCTD---LNTNNTNTTETLSIIV 115  
DB 93 NWNKNNVQEMHEDIISLWDQSLKPCVKLTPLCVTLNCTD---LNTNNTNTTETLSIIV 149  
QY 116 IIVWQRKGMKNCSEFNITTSIRDKVQREVALFYKLDVEPID--DNKN-TTNNTKYELI 173  
DB 150 -WGMERGEIKNGSEFNITTSIRDKVQREVALFYKLDVEPID--DNKN-TTNNTKYELI 205  
QY 174 NTSVITQACPKVSFEPIPIHYCTPTGFAILLKNDKFKNGTGPCTNVSTVQCTHGIRPV 233  
DB 206 NTSVITQACPKVSFEPIPIHYCAPAGFAILLKNDKFKNGTGPCTNVSTVQCTHGIRPV 265  
QY 234 VSTQLLNGSLAEVEVIRSENFNTNNAKTIIVQLNVSEINCTRENNTKRKRVTLGPGRV 293  
DB 266 VSTQLLNGSLAEVEVIRSENFNTNNAKTIIVQLNVSEINCTRENNTKRKRVTLGPGRV 325  
QY 294 WYTTGELIGNIRQAHNCISRAQWNNTLOQIATTLREQFNGK-TIAPNOSGGDPPIVMSH 352  
DB 326 WYTTGELIGNIRQAHNCISRAQWNNTLOQIATTLREQFNGK-TIAPNOSGGDPPIVMSH 385  
QY 353 FNCGEFFCYNSTOLFNSAWNTSGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGK 412  
DB 386 FNCGEFFCYNSTOLFNSAWNTSGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVGK 439  
QY 413 MYALPIKILRCSSNITGLLLTRDGGGNTDENRTHIFRPGGDMRDNRSELYKYKVKVIEP 472  
DB 440 MYALPIKILRCSSNITGLLLTRDGGGNTDENRTHIFRPGGDMRDNRSELYKYKVKVIEP 499  
QY 473 LGVAPTKARRVVQREKRAVGMFLGAGTGMGATSMALTVQARQLLSGIVQOQ 532  
DB 500 LGVAPTKARRVVQREKRAVGMFLGAGTGMGATSMALTVQARQLLSGIVQOQ 559  
QY 533 NLLRAIKAQHLLQITVWGIKQARILAVERYLKDQQLGFWGCGKLICTTAVPWN 592  
DB 560 NLLRAIKAQHLLQITVWGIKQARILAVERYLKDQQLGFWGCGKLICTTAVPWN 619  
QY 593 SWSNKTLDQIWNMTWMEWDREIDNTHLYTLIESQKQKNOQELLQDKWASL 649  
DB 620 SWSNKSLEINWMTWMEWDREIDNTHLYTLIESQKQKNOQELLQDKWASL 676

## RESULT 12

Q9DL81 PRELIMINARY; PRT; 858 AA.  
ID Q9DL81  
AC Q9DL81  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
EN Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hu, Qx, Barry A.P., Wang, Zx, Connolly S.M., Peiper S.C.,  
RA Greenberg M.L.;  
RT "Evolution of the Human Immunodeficiency Virus Type 1 Envelope during  
RT Infection Reveals Molecular Correlations of Specificity for Coreceptor  
RL Utilization and AIDS Pathogenesis."  
RL J. Virol. 74:11858-11872(2000).  
DR EMBL; AF310126; AAG37004.1; -  
DR InterPro; IPR000328; Env GP41.  
DR InterPro; IPR000777; GP120.

DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.  
SQ SEQUENCE 858 AA; 97823 MW; C0028DC1178368B CRC64;  
  
Query Match 82.2%; Score 2857; DB 15; Length 858;  
Best Local Similarity 82.5%; Pred. No. 2e-226;  
Matches 536; Conservative 44; Mismatches 56; Indels 14; Gaps 7;  
  
QY 3 LWTVYGVVPVWKEATTTLFCASDAKAYDTBAHNVWATHACVPTNPQEVLENVTNF 62  
DB 33 LWTVYGVVPVWKEATTTLFCASDAKAYDTBAHNVWATHACVPTNPQEVLENVTNF 92  
  
QY 63 NMKNVVEQMHEDIISLWDSLPKPCVKLTPLCVTLNCTDNTNTTITLSIIVWEQ 122  
DB 93 NMKNVVEQMHEDIISLWDSLPKPCVKLTPLCVTLNCTDNTNTTITLSIIVWEQ 148  
  
QY 123 RGKEMRNCSENIITTSIRDKVQREYALFKLDVEPIDDKNTNTNTK---YRLINCNTSV 179  
DB 149 EQKEIKNCSENIITGIRGKQV-EYSLFKLDVPIIDSRNNSNSTEYNSYRLINCNTSV 207  
  
QY 180 ITQACPVSPEPIPIHYCTPTGFAALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTQL 239  
DB 208 ITQACPVSPEPIPIHYCAPAGFAILKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTQL 267  
  
QY 240 LNSLAEEVIRSENFNTNAKIIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTTGE 299  
DB 268 LNSLAEEVIRSENFNTNAKIIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTTGE 326  
  
QY 300 ILGRIQAHNISRAOWNTLQOIATTLREDFGNTKIAPNOSGGDPPIVMSFNCGGEF 359  
DB 327 ILGRIQAHNISRAOWNTLQOIATTLREDFGNTKIAPNOSGGDPPIVMSFNCGGEF 386  
  
QY 360 PYCNSTOLFNSAWNTSGTWSVTRKQDGTITLPCRIKQIINRWQVGVKAMVALPIK 419  
DB 387 PYCNSTOLFNSAWNTSGTWSVTRKQDGTITLPCRIKQIINRWQVGVKAMVALPIK 441  
  
QY 420 GLIRCSNITGLLLTRDGGNGNQTEIFRPGGDMRDNWRSLEYKYKVKIEPLGVAPT 479  
DB 442 GLIRCSNITGLLLTRDGGNGNQTEIFRPGGDMRDNWRSLEYKYKVKIEPLGVAPT 501  
  
QY 480 AKRRVQREKAVGMGLGFLGAGSTWGTSMALTVOARQLLSGIVQOQNLLRAI 539  
DB 502 AKRRVQREKAVGMGLGFLGAGSTWGTSMALTVOARQLLSGIVQOQNLLRAI 561  
  
QY 540 KAOQHLLQLTWGIKQOLARILAVERYLKDQGLGFCGSGKLICTTAVPNWASNSKTL 599  
DB 562 KAOQHLLQLTWGIKQOLARILAVERYLKDQGLGFCGSGKLICTTAVPNWASNSKTL 621  
  
QY 600 DQIWNMTWMEWDREIDNTHLYTLIESQNQOEKNOQELLQDKWASL 649  
DB 622 DQIWNMTWMEWDREIDNTHLYTLIESQNQOEKNOQELLQDKWASL 671

RESULT 13  
Q9J022 PRELIMINARY; PRT; 853 AA.

AC Q9J022;  
DT 01-OCT-2000 (TremBLrel. 15, Created)  
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TremBLrel. 22, Last annotation update)  
DE Envelope glycoprotein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OC NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BORI;  
RX MEDLINE=20091324; PubMed=10623731;  
RA Shieh J.T., Martin J., Baltuch G., Malim M.H., Gonzalez-Scarano F.,  
RT "Determinants of syncytium formation in microglia by human  
RT immunodeficiency virus type 1: role of the V1/V2 domains."

RL J. Virol. 74:693-701 (2000).  
DR EMBL; AF227273; AAF61104.1; -  
DR InterPro; IPR000328; Env\_GP41.  
DR Pfam; PF00516; GP120.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.  
SQ SEQUENCE 853 AA; 96346 MW; 1601A1A72468E377 CRC64;  
  
Query Match 82.2%; Score 2855.5; DB 15; Length 853;  
Best Local Similarity 82.8%; Pred. No. 2.7e-226;  
Matches 538; Conservative 41; Mismatches 54; Indels 17; Gaps 7;  
  
QY 2 NLWTVYGVVPVWKEATTTLFCASDAKAYDTBAHNVWATHACVPTNPQEVLENVTNF 61  
DB 32 NLWTVYGVVPVWKEATTTLFCASDAKAYDTBAHNVWATHACVPTNPQEVLENVTNF 91  
  
QY 62 FMWKNVVEQMHEDIISLWDSLPKPCVKLTPLCVTLNCTDNTNTTITLSIIVWE 121  
DB 92 FMWKNVVEQMHEDIISLWDSLPKPCVKLTPLCVTLNCTDNTNTTITLSIIVWE 145  
  
QY 122 QRGKEMRNCSENIITTSIRDKVQREYALFKLDVEPIDDKNTNTNTKYRLINCNTSVIT 181  
DB 146 QRGKEMRNCSENIITTSIRDKVQREYALFKLDVEPIDDKNTNTNTKYRLINCNTSVIT 200  
  
QY 182 QACPKVSPEPIPIHYCTPTGFAALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTQLLN 241  
DB 201 QACPKVSPEPIPIHYCTPTGFAALLKNDKFKNGTGPCTNVSTVQCTHGIRPVVSTQLLN 260  
  
QY 242 GSLAEEVIRSENFNTNAKIIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTTGEIL 301  
DB 261 GSLAEEVIRSENFNTNAKIIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTTGEIL 320  
  
QY 302 GNIRQAHNISRAOWNTLQOIATTLREDFGNTKIAPNOSGGDPPIVMSFNCGGEF 360  
DB 321 GNIRQAHNISRAOWNTLQOIATTLREDFGNTKIAPNOSGGDPPIVMSFNCGGEF 380  
  
QY 361 YCNSTOLFNSAWNTSGTWSVTRKQDGTITLPCRIKQIINRWQVGVKAMVALPIK 420  
DB 381 YCNSTOLFNSAWNTSGTWSVTRKQDGTITLPCRIKQIINRWQVGVKAMVALPIK 436  
  
QY 421 LIRCSNITGLLLTRDGGNGNQTEIFRPGGDMRDNWRSLEYKYKVKIEPLGVAPT 479  
DB 437 LIRCSNITGLLLTRDGGNGNQTEIFRPGGDMRDNWRSLEYKYKVKIEPLGVAPT 496  
  
QY 480 AKRRVQREKAVGMGLGFLGAGSTWGTSMALTVOARQLLSGIVQOQNLLRAI 539  
DB 497 AKRRVQREKAVGMGLGFLGAGSTWGTSMALTVOARQLLSGIVQOQNLLRAI 556  
  
QY 540 KAOQHLLQLTWGIKQOLARILAVERYLKDQGLGFCGSGKLICTTAVPNWASNSKTL 599  
DB 557 KAOQHLLQLTWGIKQOLARILAVERYLKDQGLGFCGSGKLICTTAVPNWASNSKTL 616  
  
QY 600 DQIWNMTWMEWDREIDNTHLYTLIESQNQOEKNOQELLQDKWASL 649  
DB 617 DQIWNMTWMEWDREIDNTHLYTLIESQNQOEKNOQELLQDKWASL 666

RESULT 14

Q9E1R7 PRELIMINARY; PRT; 855 AA.  
ID Q9E1R7  
AC Q9E1R7;  
DT 01-MAR-2001 (TremBLrel. 16, Created)  
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)  
DT 01-OCT-2002 (TremBLrel. 22, Last annotation update)  
DE Envelope protein.  
GN ENV.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OC NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20442410; PubMed=10984542;

```

RA Cleghorn F.R., Carr J.K., Edwards J., Mahabir B., Sill A.,
RA McDaniel C.B., Connolly S.M., Goodman D., Bennetts R.Q., O'Brien T.R.,
RA Weinhold K.J., Bartholomew C., Blattner W.A., Greenberg M.L.;
RT "A distinctive clade B HIV type 1 is heterosexually transmitted in
RT Trinidad and Tobago.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:10532-10537(2000).
DR EMBL; AF277071; AAG22514.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 855 AA; 97102 MW; 26271D6CCCCCFAC CRC64;

Query Match      82.2%; Score 2855.5; DB 15; Length 855;
Best Local Similarity 82.0%; Pred. No. 2.7e-226;
Matches 534; Conservative 42; Mismatches 52; Indels 23; Gaps 6;

QY 4 WTVTVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQVVLNVNTENF 63
DB WTVTVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQVVLNVNTENF 95
QY 64 MKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNC-----TDLNTNTNTTTELSIV 118
DB MKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNC-----TDLNTNTNTTTELSIV 118
QY 96 MKNNVVDQMHEDIISLWQSLKPCVKLTPLCVTLNCIDYVKNNDTNNTNNTNNSVME 155
DB MKNNVVDQMHEDIISLWQSLKPCVKLTPLCVTLNCIDYVKNNDTNNTNNTNNSVME 155
QY 119 VMEQRKGEMRNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNNTKYLNCNTS 178
DB VMEQRKGEMRNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNNTKYLNCNTS 178
QY 156 T-----GEMKNCSPNITTSVRDKVREYALFYKLDVTPIDDN---NDNSGYRLNCNTS 206
DB T-----GEMKNCSPNITTSVRDKVREYALFYKLDVTPIDDN---NDNSGYRLNCNTS 206
QY 179 VITQACPKVSEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 238
DB VITQACPKVSEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 238
QY 207 VITQACPKVSEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 266
DB VITQACPKVSEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 266
QY 239 LINGSLAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 298
DB LINGSLAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 298
QY 267 LINGSLAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 326
DB LINGSLAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 326
QY 299 EILGNIRQAHNCISRAQWNNTLQOIATTLREQFNKTIAPNOSGGDPPIVHVSFNCGE 358
DB EILGNIRQAHNCISRAQWNNTLQOIATTLREQFNKTIAPNOSGGDPPIVHVSFNCGE 358
QY 327 EILGNIRQAHNCISRAQWNNTLQOIATTLREQFNKTIAPNOSGGDPPIVHVSFNCGE 386
DB EILGNIRQAHNCISRAQWNNTLQOIATTLREQFNKTIAPNOSGGDPPIVHVSFNCGE 386
QY 359 FYCNSITQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVKAMYPALPI 418
DB FYCNSITQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVKAMYPALPI 418
QY 387 FFYCNSTQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVKAMYPALPI 438
DB FFYCNSTQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVKAMYPALPI 438
QY 419 KGLIRCSSNITGLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIPLGVAPT 478
DB KGLIRCSSNITGLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIPLGVAPT 478
QY 439 RGOISCSNITGLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIPLGVAPT 498
DB RGOISCSNITGLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIPLGVAPT 498
QY 479 KAKRRVQREKRAVGMGLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538
DB KAKRRVQREKRAVGMGLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 538
QY 499 KARRVQREKRAVGMGLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 557
DB KARRVQREKRAVGMGLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 557
QY 539 IKAQOHLQLTVMGIIKQOQARILAVERYLKQOQLLGFWSGKLICTTAVPWNASNKT 598
DB IKAQOHLQLTVMGIIKQOQARILAVERYLKQOQLLGFWSGKLICTTAVPWNASNKT 598
QY 558 IEAQOHLQLTVMGIIKQOQARILAVERYLKQOQLLGFWSGKLICTTAVPWNASNKT 617
DB IEAQOHLQLTVMGIIKQOQARILAVERYLKQOQLLGFWSGKLICTTAVPWNASNKT 617
QY 599 LDOIWNNTWMDREIDNTYLIYLTILBESQOQKQOELLQDLKWSL 649
DB LDOIWNNTWMDREIDNTYLIYLTILBESQOQKQOELLQDLKWSL 649
QY 618 MDDIWNNTWMDREIDNTYLIYLTILBESQOQKQOELLQDLKWSL 658
DB MDDIWNNTWMDREIDNTYLIYLTILBESQOQKQOELLQDLKWSL 658

RESULT 15
Q9DH32 PRELIMINARY; PRT; 858 AA.
ID Q9DH32
AC Q9DH32;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein (Envelope protein).
GN ENV.
OS Human immunodeficiency virus 1.

```

```

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_taxonomy:116176;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=QH1420C.2;
RX MEDLINE=20541986; PubMed=11090186;
RA Hu, Qx, Barry A.P., Wang, Zx, Connolly S.M., Peiper S.C.,
RA Greenberg M.L.;
RA "Evolution of the Human Immunodeficiency Virus Type 1 Envelope during
RT Infection Reveals Molecular Corollaries of Specificity for Coreceptor
RT Utilization and AIDS Pathogenesis.";
RL J. Virol. 74:11858-11872(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20442410; PubMed=10984542;
RA Cleghorn F.R., Carr J.K., Edwards J., Mahabir B., Sill A.,
RA McDaniel C.B., Connolly S.M., Goodman D., Bennetts R.Q., O'Brien T.R.,
RA Weinhold K.J., Bartholomew C., Blattner W.A., Greenberg M.L.;
RT "A distinctive clade B HIV type 1 is heterosexually transmitted in
RT Trinidad and Tobago.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:10532-10537(2000).
DR EMBL; AF310124; AAG37002.1; -.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 858 AA; 97843 MW; 6B77091DC964BF94 CRC64;

Query Match      82.2%; Score 2854; DB 15; Length 858;
Best Local Similarity 82.6%; Pred. No. 3.6e-226;
Matches 537; Conservative 42; Mismatches 57; Indels 14; Gaps 7;

QY 3 LWTVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQVVLNVNTENF 62
DB LWTVYGVVWKEATTTILFCASDAKAYDEAHNVWATHACVPTNPQVVLNVNTENF 92
QY 63 NWKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCIDYVKNNDTNNTTTELSIVVMEQ 122
DB NWKNNVVEQMHEDIISLWQSLKPCVKLTPLCVTLNCIDYVKNNDTNNTTTELSIVVMEQ 122
QY 93 NWKNNVVDQMHEDIISLWQSLKPCVKLTPLCVTLNCIDYVKNNDTNNTTTELSIVVMEQ 148
DB NWKNNVVDQMHEDIISLWQSLKPCVKLTPLCVTLNCIDYVKNNDTNNTTTELSIVVMEQ 148
QY 123 RKGEMRNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNNTK---YRLNCNTSV 179
DB RKGEMRNCSFNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNNTK---YRLNCNTSV 179
QY 149 EKGGEIKNCSPNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNNTK---YRLNCNTSV 207
DB EKGGEIKNCSPNITTSIRDKVQREYALFYKLDVEPIDDKNTNTNNTK---YRLNCNTSV 207
QY 180 ITQACPKVSEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 239
DB ITQACPKVSEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 239
QY 208 ITQACPKVSEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 267
DB ITQACPKVSEPIPIHYCTPTGTFALLKNDKFKNGTGTCTNVSTVQCTHGIRPVVSTQL 267
QY 240 LINGSLAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 299
DB LINGSLAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 299
QY 268 LINGSLAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 326
DB LINGSLAEEVWIRSENFTNNAKTIIVQLNVSEINCTRPNNHTRKRVTLGPRVWYTG 326
QY 300 ILGNIRQAHNCISRAQWNNTLQOIATTLREQFNKTIAPNOSGGDPPIVHVSFNCGE 359
DB ILGNIRQAHNCISRAQWNNTLQOIATTLREQFNKTIAPNOSGGDPPIVHVSFNCGE 359
QY 327 IIGDIRQAHNCISRAQWNNTLQOIATTLREQFNKTIAPNOSGGDPPIVHVSFNCGE 386
DB IIGDIRQAHNCISRAQWNNTLQOIATTLREQFNKTIAPNOSGGDPPIVHVSFNCGE 386
QY 360 FYCNSITQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVKAMYPALPI 419
DB FYCNSITQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVKAMYPALPI 419
QY 387 FYCNSITQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVKAMYPALPI 441
DB FYCNSITQLFNSAWNTSNGTWSVTRKQKDTGDIITLPCRIKQIINRWQVVKAMYPALPI 441
QY 420 GLIRCSSNITGLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIPLGVAPT 479
DB GLIRCSSNITGLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIPLGVAPT 479
QY 442 GQIRCSSNITGLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIPLGVAPT 501
DB GQIRCSSNITGLLTRDGGGNGTTEIFPPGGGDMRDNRSELYKYKVKIPLGVAPT 501
QY 480 AKRRVQREKRAVGMGLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 539
DB AKRRVQREKRAVGMGLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 539
QY 502 AKRRVQREKRAVGMGLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 561
DB AKRRVQREKRAVGMGLGFLGAAGSTMGATSMALTVOARQLLSGIVQOQNLLRA 561
QY 540 KAOQHLQLTVMGIIKQOQARILAVERYLKQOQLLGFWSGKLICTTAVPWNASNKT 599
DB KAOQHLQLTVMGIIKQOQARILAVERYLKQOQLLGFWSGKLICTTAVPWNASNKT 599
QY 562 EAOQHLQLTVMGIIKQOQARILAVERYLKQOQLLGFWSGKLICTTAVPWNASNKT 621
DB EAOQHLQLTVMGIIKQOQARILAVERYLKQOQLLGFWSGKLICTTAVPWNASNKT 621

```

Qy 600 DQIWNNTWMEWDREIDNVTHTLITLIESONQOEKNOELLOLDKWASL 649  
Dd 622 TDIWNTWMOQDREITNTYELIYTLLEDAQNOQEKNQODLLELDKWASL 671

Search completed: January 13, 2004, 09:11:26  
Job time : 44 secs